

Gene donors' rights at risk

Even those who treat other people's genomes as intellectual property concede that sample donors deserve some share of their genes' financial value. But those rights are already threatened in several ways — a situation that requires urgent attention.

FEW of the ethical dilemmas created by the growing commercial interest in human genetics are as complex — or as potentially explosive — as those relating to the 'ownership' of genetic data. So far, the most acute conflicts in this area have been provoked by applications from Western researchers for patents on genetic information derived from indigenous groups in developing countries. This is not surprising. Such instances stimulate passionate feelings about the way other resources belonging to such groups have been exploited in the past, and how little of the fruits of this exploitation made their way back. But it would be wrong either to dismiss the passion as being purely politically motivated, or to see the issue as restricted to indigenous populations.

Virtually all large pharmaceutical industries are now developing an intense interest in the genetic basis of disease, which many see as central to their future products. The increasing sophistication of the tools of genetic analysis means that this interest is no longer restricted to 'single-gene' diseases such as cystic fibrosis, but now extends to a wide spectrum of what are called 'complex disorders', from arthritis to cancer. Companies likely to achieve dominance are those that maintain a scientific lead over their competitors. And this is likely to be built partly on a combination of commercial secrecy and the exploitation of intellectual property.

All this means that the stakes involved in the quest for genetic information are high. Those already caught up in this new gold-rush include virtually any group considered sufficiently homogeneous to provide tissue or blood samples from which information leading to the eventual identification of a disease-related gene or genes can be identified (see Briefing, page 12). Such groups range from isolated populations distinguished by certain medical characteristics (which could be resistance to a particular disease) to sets of families suffering from a common illness, such as diabetes, who agree to assist researchers in the hope of developing improved treatment.

The dilemma raised can be expressed straightforwardly: if the knowledge about the gene or genes in question subsequently attains a significant commercial value, how much of that value should be returned to the group that provided the original samples, and in what form? In the past, the prospect of novel treatment resulting from a better understanding of disease has generally been seen as sufficient reward. This formula is no longer sufficient, chiefly because it does not provide any mechanism for ensuring that other potential benefits are made available to contributors. But finding a substitute mechanism for distributing the gains equitably is fraught with practical difficulties.

For example, the proposal that individuals contributing samples to a research project should each receive a slice of future royalty payments sounds attractive in principle. But imagine trying to negotiate such payment on behalf of 5,000 families, spread over three continents, involved in a study of a particular gene. There are also legitimate philosophical questions. The implications of part of an individual's genetic identity being something that can be bought and sold — that is, reduced to a commodity — flies in the face of many deeply held cultural beliefs about human value, even if it is already enshrined *de facto* in Western patent laws.

Three factors need to be recognized by all those engaged in such debates. First, informed-consent agreements should explicitly

acknowledge the possibility of commercial gains from research, and indicate how the issue is to be equitably addressed. Second, any agreement with commercial sponsors should not obstruct the development of research, for example by imposing excessively strict limitations on who should have access to family data, and under what conditions; legislation that already permits this should be re-examined to see whether the rights given to patent holders are so broad as to discourage healthy competition. Third, cultural diversity, as well as the political rights of indigenous populations, must be respected in international agreements on patent law — perhaps by introducing clauses that allow for both in the patenting of human genetic material in future revisions of GATT, the General Agreement of Trade and Tariffs. □

Budget clouds lift a little

Last week's US budget settlement leaves science in a stronger position than was feared, but this is no time to relax.

THE United States budget for the 1996 financial year has now been fixed, after seven months of unseemly haggling between the President and the Congress. It is a powerful tribute to the political skills of President Bill Clinton that he is widely perceived as emerging from this process as the victor. For its outcome is an unprecedented \$22 billion cut in non-defence spending at the discretion of the Congress. That is a 9 per cent reduction in one year — 12 per cent if inflation is allowed for.

Given such a context, science has not fared too badly (see pages 6–7). Non-defence research and development (R&D) spending has fallen, but technology, not science, has taken most of the cuts. The largest basic research agency, the National Institutes of Health, will enjoy a 6 per cent budget increase. The National Science Foundation and the National Aeronautics and Space Administration have cut back on money for new facilities, while their support for research has been frozen. Non-defence R&D at the Department of Energy is down by 10 per cent but most of the attrition has been in technology projects, not in pure science.

The picture is worse for small agencies, with the National Biological Survey and various activities related to global climate change falling victim to political micromanagement of the worst kind. But overall, the science lobby has done reasonably well. It is those without a lobby, such as public housing residents and people in need of legal aid, who bear the real brunt of this budget.

Now attention turns to 1997 and beyond. The Clinton administration helpfully suggests that we ignore its projections for 1998 onwards (see *Nature* 380, 657; 1996) although it cannot openly state the reason why: whether Clinton wins or loses in November, they will be history by Christmas. The Republicans in Congress are also clicking into election mode. In preparing their 1997 budget, the early signs are that they will talk less about spending cuts and more about tax cuts. The glory days of science funding may be gone for ever, but this year's budgetary environment should be somewhat gentler than that of the one just ended. Nevertheless, the need to lobby hard on all fronts persists. □



Plans for Russian uranium payments stir fears of nuclear proliferation

Paris. Russia may provide highly enriched uranium (HEU) to two research reactors in France as its subscription fee to research programmes at the two centres, according to an agreement on nuclear materials reached last month by France and the Russian ministry of atomic energy.

But groups supporting nonproliferation claim that the plan flies in the face of international efforts to reduce civilian commerce in weapons-grade nuclear material, and would increase the risk of such material falling into terrorist hands.

Final details have yet to be agreed. But as part of the agreement, Russia would supply HEU to the Orphée reactor, which is owned by the French Atomic Energy Commission (CEA), and the 58.3-MW reactor at the Institut Laue-Langevin (ILL) in Grenoble — the world's most powerful neutron source (see *Nature* 379, 284; 1996).

ILL is owned by France and Germany — each pay annual subscriptions of around FF100 million (US\$20 million) — and the United Kingdom, which pays about FF67 million. The reactor also has three 'scientific members', Austria, Spain and Switzerland, which contribute a total of around FF20 million annually.

Under the proposed agreement, Russia would also be given the status of scientific member, according to Reinhard Scherm, director of ILL, who says that Russian membership will provide an important scientific boost to ILL's activities. Russia would pay its fee in HEU; the precise amount of HEU involved in the deal has not been disclosed, but ILL consumes around 45 kg a year.

ILL has been seeking new partners since the United Kingdom reduced its annual subscription from a third to a quarter in 1992, forcing ILL to reduce the number of its beamlines and to make 80 of its 500 staff redundant. It is now actively wooing Italy, according to Scherm.

But the prospect that Russia will pay for its membership in HEU has provoked fierce protests. Paul Leventhal, president of the Nuclear Control Institute (NCI) — a lobby group based in Washington DC — describes the proposal as a "direct assault" on the 1978 international agreement on Reduced Enrichment for Research and Test Reactors (RERTR). This is aimed at phasing out the use of HEU in research reactors, given the danger that the fuel could be used to make nuclear weapons.

The United States has reinforced the RERTR programme by banning the export of HEU to reactors that could be converted

to light-enriched (non-weapons grade) uranium (LEU) but have declined to do so, and also allowing exports to other reactors only on condition that they convert to LEU as soon as is feasible (see *Nature* 369, 89; 1994). This policy has led to a reduction in annual US exports of HEU from more than a tonne in the 1970s to zero.

As a result, nearly all research reactors in Europe and the United States have either converted to LEU or are in the process of doing so. A handful of reactors that could use existing LEU fuels have refused to convert, however, and about a dozen have not converted because suitable LEU fuel substitutes have not yet been developed.

Leventhal argues that if Russia is allowed to provide Europe with HEU, the US ban on HEU exports would no longer be an effective incentive to force research reactors to convert to LEU — until now, the United States had been the almost exclusive supplier of HEU. In a bid to discourage Russia from exporting HEU, US agencies have also begun buying up HEU extracted from Russian warheads and diluting it to LEU.

Russian exports of HEU to Europe

would undermine many of the gains made under the RERTR programme, says Leventhal. Such exports would, for example, enable construction of the controversial FRM-II research reactor outside Munich in Germany (see *Nature* 379, 284; 1996), which would be the first HEU-fuelled research reactor to be built — outside Libya and China — since the establishment of the RERTR programme. And Russian commerce in HEU would defeat longer-term plans to encourage research reactors in Russia and China to convert to LEU.

Attempts to persuade Germany not to use HEU in the FRM-II would be "compromised" if France imports Russian HEU, argues Mycle Schneider of the Paris-based energy and environmental consultancy WISE. "The French deal sets a very bad precedent that goes against all attempts to get rid of civil HEU," he says.

Indeed, Russia is already negotiating the supply of HEU to research reactors in the United Kingdom, Germany, Belgium and the Netherlands, according to Euratom (the European Atomic Energy Community), although formal agreements have not yet ►

UK observatories look to private sector

London. The British government has announced that private-sector organizations will be invited to bid for the services provided by its 'Royal Observatories', made up of telescopes on Hawaii and the Canary Islands, and instrumentation development and control activities at the Royal Greenwich Observatory in Cambridge and the Royal Observatory Edinburgh.

The intention is to make more efficient use of the observatories' annual budget of £15 million (US\$10 million), in line with recommendations put forward by the Particle Physics and Astronomy Research Council (PPARC), their legal owner.

PPARC has welcomed the announcement and intends to use any savings or additional income from potential new observatory activity to help fund long-term scientific projects.

But the decision has been criticized by at least one major trade union. Tony

Bell, a negotiator with the Institute of Professionals, Managers and Specialists (IPMS), says that private-sector management of the observatories is not the solution to what he describes as "inadequate funding" for PPARC.

He warns that the move will sound "the death knell" for the Edinburgh and Cambridge organizations, both of which support the Isaac Newton Group of telescopes at La Palma in the Canary Islands (above), and the Joint Astronomy Centre on Hawaii. Ehsan Masood

► been reached. Any such agreement would need to be signed by Euratom, which is responsible for enforcing safeguards on all civil nuclear materials within Europe.

Similarly, Leventhal describes the proposed deal between France and Russia as evidence that France feels immune from international norms on nuclear nonproliferation. "It's just a further attempt by the French to stick a thumb in the eye of the US in the nonproliferation field," he claims.

The United States has applied strong diplomatic pressure to Euratom and Russia to prevent them from opening up trade in HEU. Indeed, some US congressmen had pressed for the United States to block renewal of the US-Euratom agreement — which covers trade in nuclear materials — until Europe and Russia agree not to trade in HEU (see *Nature* 379, 760; 1996).

While such diplomatic pressure is continuing, the signing of the US-Euratom agreement itself represents a setback for US efforts. The agreement contains a letter from Stuart Eisenstat, the US ambassador to Euratom, which states that the United States is "committed to eliminating [the civil use of HEU] over time".

But the letter also states that the United States recognizes "that specific research reactors in the European Atomic Energy Community may, under certain circumstances, need to use highly-enriched uranium as fuel". This loophole is reported to have been included after intense European pressure.

Costas Verros, a spokesman for Euratom, says that the letter is explicit recognition that several European reactors cannot convert to LEU at present. "The matter is settled," says Verros, adding that Euratom has "an obligation to help operators to find the fuel they need". Verros points out that the European Union has signed an agreement with Russia similar to that between Euratom and the United States on the peaceful uses of nuclear technology, and that this allows European countries to seek HEU in Russia.

Declan Butler

Drug company 'suppressed' publication of research

Washington. Controversy over the secrecy demanded of biomedical researchers by pharmaceutical companies captured the limelight again last week, with a report that a company had suppressed the publication of a journal article out of fear that the conclusions would hurt the multi-million dollar sales of one of its products, a thyroid drug.

The story, first reported in *The Wall Street Journal*, tells how the British company Boots Co. apparently succeeded in persuading Betty Dong, a researcher it had funded at the University of California, San Francisco (UCSF), to withdraw a paper scheduled to appear in the *Journal of the American Medical Association (JAMA)*.

Soon after the article was due to have appeared in January 1995, Boots' drug division was bought by Germany's BASF AG, for \$1.4 billion. Synthroid, a drug made by the division which is used in hypothyroidism, and which accounts for 84 per cent of the \$600-million US market for thyroid replacement drugs, is thought to have figured prominently in the price paid.

Boots' huge share of the market reflects the inability of its rivals to prove beyond doubt that their products were 'bioequivalent' to Synthroid. In the 1980s, Boots' drug division — now part of Knoll Pharmaceutical Co., a New Jersey division of BASF — decided to prove its rivals' inferiority for good. The company sought out Dong, a clinical pharmacist, and paid her \$250,000 to carry out a comparative study of Synthroid and three alternative drugs.

But the results were not what Boots expected. Dong and her research team found that the alternatives were 'bioequivalent' — that they were absorbed in the blood in the same way as Synthroid — and that use of the significantly cheaper, equally effective

alternatives would reduce US health-care costs by \$356 million a year.

After several years spent trying to discredit Dong's findings — including hiring private investigators to search her background for conflicts of interest — Boots succeeded in blocking publication, using the fact (which neither side denies) that when Dong began the work, she signed a contract promising the results would not be published "without written consent" of the company. The company allegedly threatened Dong and her colleagues with a lawsuit if she published — an allegation that a Boots executive adamantly denies.

Faced with the prospect of ruinous legal fees, as well as a reversal by UCSF, which initially backed her efforts to publish, but then said it could not because of the legal risk, Dong told *JAMA* to cancel her paper, which was already at the printer. She has not succeeded in finding another publisher.

Meanwhile, Gilbert Mayor, formerly director of medical services for Boots and now the senior director of medical research at Knoll, was the lead author on a lengthy critique of her unpublished findings in the *American Journal of Therapeutics*, of which Mayor is an editor, in June 1995.

Carter Ekert, the Boots executive who instigated Dong's study, told *The Wall Street Journal*: "I stopped a flawed study that would have put millions of patients at risk." Notwithstanding its acceptance by *JAMA* peer reviewers, the company has argued that the study contained mistakes in data analysis and patient management, among other things, due to factors which the reviewers would not have been aware of.

UCSF officials say that the study is "valid" and without "significant" flaws. They add that Dong violated university policy by signing the contract, calling its limitation on her right to publish "not acceptable".

The university also defends its decision not to make its legal resources available to Dong and her colleagues. In its statement, it calls the right to publish "the essence of what a research institution is all about". But, it continues, "the difficulty here is weighing the right to publish against a likely claim against the University for breach of contract and the possibility of significant damages".

Dong told *The Wall Street Journal* that signing the contract was naïve, that she ought not to have signed, and that she will not do drug company-sponsored research again. Stephen Rosenberg, the chief of surgery at the National Cancer Institute in Bethesda, Maryland, and a strong critic of biomedical secrecy, says of the episode: "I'm shocked. But I'm not surprised."

Meredith Wadman

Austrian academics fight teaching fee review

Munich. The Austrian research minister, Rudolf Scholten, has invited university academics and students, who have been on strike since Easter, to suggest how the country should save ÖS1 billion (US\$94 million) from the university budget over the next two years.

Scholten's own proposals for making the large savings dictated by Austria's financial crisis had been greeted with horror by students, who could lose privileges such as free public transport, and by academic staff, particularly the so-called *Mittelbau*, equivalent to assistant and associate professors, who could lose a large proportion of the generous teaching fees on which they rely

because most are employed by universities only part-time.

Scholten wants to save nearly half of his ÖS1 billion target by paying low fees — or no fees at all — to those who give a few lectures a term, and then pay progressively higher fees for more teaching hours. This is opposed by groups such as the Federal Conference for Scientific Staff, which argues that this system would encourage young academics to neglect research in favour of doing more teaching to maintain their income.

But Scholten argues that the quality of Austrian research has already suffered from the current system of piece-work.

Quirin Schliermeier

Savings drive unsettles Australian science

Sydney. Australian researchers are facing four months of uncertainty before they know how much their research programmes will suffer as part of the new coalition government's general drive to save \$A8 billion (US\$6.7 billion) over its first two budgets. The government has already started to make big cuts in all departments, and direct funding for science and technology — as well as indirect support through universities — is coming under close scrutiny.

Eight weeks after John Howard became prime minister on the defeat of the Labor government in March's general elections, the first of three rounds aimed at reducing the size of the public service has started. About 200 staff will leave the Department of Industry, Science and Tourism by July, and nearly 1,800 will leave the Department of Employment, Education, Training and Youth Affairs which oversees universities.

But the resulting savings will not be sufficient to meet the government's target, and cuts to research programmes will also be necessary. The first indication of shrinkage in the previous government's research and development programme came when a proponent of a High Performance Computing and Communications programme revealed that the \$A30 million allocated "has already been halved and funds are evaporating fast".

Peter McGauran, the new science and technology minister, is already having talks with research organizations, which say that their concerns are being listened to. But any scientist whose research is supported by the government is urgently seeking confirmation of rumoured cuts.

McGauran confirms that pre-election promises for science and technology will be met (see *Nature* 380, 192; 1996). But he points out that all areas of government are being examined for savings based on "duplication or lack of priority", and that science and technology "are no different". He has not ruled out using cuts in base funding to meet earlier promises; the final results will be known when the government announces its budget plans on 20 August.

Among initiatives receiving the closest scrutiny are several only recently announced by the previous government. Seven Major National Research Facilities and three 'Innovation Flagships' announced in Labor's Innovation Statement (see *Nature* 378, 653; 1995) appear to have been secured with contracts. But the funding of other new programmes, such as the computing project, is vulnerable.

Also in serious doubt is a move by Australian astronomers to become paid-up participants in the world's largest optical telescope being built in Chile by the European Southern Observatory, which was supported by the former Science Minister, Peter Cook, shortly before the 2 March election

(see *Nature* 379, 668; 1996).

Particular fears are being expressed in universities following a bruising first meeting between the Australian Vice-Chancellors' Committee (AVCC) and Amanda Vanstone, the new Education Minister. Fay Gale, vice-chancellor of the University of



McGauran: job losses at CSIRO 'inevitable'.

Western Australia and president of the AVCC, has warned that budget reductions of 5 to 10 per cent are likely to be forced on universities, and has asked each university to specify what it would have to cut.

The AVCC, having previously welcomed the Coalition's pre-election promise to "maintain the levels of funding to universities in terms of operating grants", is now sceptical. Gale said reneging on commitments "would have a disastrous effect on university teaching and research".

The restructuring of Australia's major science agency, the Commonwealth Scientific and Industrial Research Organisation, by removing the upper echelon of six institutes (see *Nature* 380, 276; 1996), may help the government to direct more money to

research. Resolving the continuing problems of the troubled Division of Animal Production may offer immediate savings.

Despite pressure from his rural constituency, McGauran accepts the need for "inevitable staff redundancies" in this division, and the sale of pastoral properties to reduce its crippling debt of more than \$A5 million and budget overruns due to reduced funding from the wool industry. He has asked for options to be presented for a "countercyclical approach" to funding the division, aimed at insulating it from the short-term fluctuations in commodity prices.

Following a rocky relationship with the previous government, CSIRO appears to have made a constructive start with the new government. "We are pleased with the active interest of both ministers in the organisation, and with their commitment to delivering the promised increase of A\$20 million per annum for three years on our existing base funding," says Adrienne Clarke, chair of the CSIRO Board.

McGauran has moved to reassure the 62 Cooperative Research Centres by affirming the government's "strong commitment to maintaining" the CRC programme, and by approving five more CRCs. However, a new committee to oversee the programme will be asked "to test the centres very rigorously against their agreed outcomes and the CRC evaluation criteria".

Peter Pockley

AIDS office opposes outside planning

Washington. Senior research managers at the US National Institutes of Health (NIH) have taken issue with the recommendation in a recent report on federally funded AIDS research, arguing that the direction of such research should remain firmly in the hands of institute directors and their advisers.

The report, issued in March, had suggested that this role should, in part, be assumed by the panels of external scientists responsible for reviewing research grants. It said that members of these panels should be involved in setting AIDS research priorities, and that the panels should judge competing applications in part on the extent to which they meet such priorities — which are yet to be decided.

At present, peer review groups, or 'study sections', judge proposals strictly on scientific merit, not on the nature of the research. But the report of a panel of 118 scientists and others, led by Arnold Levine, a molecular biologist at Princeton University, New Jersey, and published in April, suggests that external review panels should be "better informed" of scientific priorities for AIDS research, and should "consider these priorities in their review".

According to William Paul, director of

the NIH Office of AIDS Research, this recommendation has not met with enthusiasm among directors of NIH's 24 institutes, divisions and centres. He told a congressional subcommittee last week that there is 'universal' opposition among the directors to involving the grant-reviewing scientists in setting research priorities.

"We felt the [Levine] panel had probably not given adequate consideration" to the matter, Paul told John Porter (Republican, Illinois), the chairman of subcommittee of the House of Representatives appropriations committee that funds the NIH. "It is essential that the peer review process be removed from the planning process," Paul said. "The purpose of peer review is to make the scientific judgement as to whether or not a grant is meritorious in its own right."

Harold Varmus, the director of NIH, agrees. "This is not to say that the significance of the research is not an issue in the initial peer review process," Varmus told the same subcommittee hearing. "But to ask the peer reviewers to try to align their evaluation [with the] priorities of various institutes and the Office of AIDS Research to my mind would be a mistake."

Meredith Wadman

Japanese plan for doubling of public science spending

Tokyo. Japan's public sector spending on science and technology will nearly double within four years if a resolution endorsed by a committee within the ruling coalition's strongest party, the Liberal Democratic Party (LDP), is adopted by the government. Implementation of the new resolution would fulfil a promise made in 1992 to double science spending by the turn of the century.

The resolution, agreed at a meeting of a recently formed LDP advisory committee on science and technology, calls for the national government to increase annual spending on science and technology to ¥4,660 billion (US\$43.7 billion) by the end of the decade, 1.7 times current spending. Under the plan, research funds will double to ¥2,460 billion, funds for personnel development and exchanges will quadruple to ¥120 billion, and funds for basic infrastructure and maintenance treble to ¥585 billion.

An LDP spokesman says a joint declaration formally adopting the resolution is expected from all the coalition's parties next month. The committee that passed this resolution has also been charged with advising the party's executive on framing the Basic Science and Technology Plan, formulation of which is required under a new science and technology law passed late last year (see *Nature* 378, 227; 1995).

Omi Koji, the LDP member responsible for drafting the law, is also playing a leading role in the new committee. He and other LDP politicians have managed to increase funding for science and technology despite the continuing recession and the government's commitment to cover the enormous losses of failed financial institutions.

During the drafting of the new plan for science, the committee heard from Susumu Tonegawa, joint winner of the 1987 Nobel prize for physiology or medicine. Tonegawa's testimony included stinging criticism of the university *koza*, or professorial chair system. The rigid hierarchical nature of this system stifles creativity in young researchers and is why Japan has obtained relatively few Nobel prizes, Tonegawa is reported to have said.

The resolutions for radical increases in funding have political support within the coalition, as well as the backing of ministries and agencies which will benefit from larger budgets. But Japan's Ministry of Finance, whose approval must be sought for all expenditure, is said to be resisting the new plans. **Stephen Barker**

US budget deal frees money

Washington. President Bill Clinton and the US Congress have finally agreed a firm budget for the 1996 financial year — which started last October — ending seven months of chaos and uncertainty at science funding agencies, as in other government departments.

In a deal finalized late last week which finally broke the budget impasse, the administration managed to eliminate several proposed relaxations of environmental law that Congress had sought to attach to the budget.

The two sides also reached a compromise on some of Clinton's favourite programmes. These included the Advanced Technology Program (ATP), which Congress Republicans have wanted to shut down. It will now receive \$221 million, compared with the \$491 million Clinton originally requested.

One surprise provision in the budget deal will prevent redundancies at the headquarters of the National Aeronautics and Space Administration announced only two weeks ago (see *Nature* 380, 657; 1996).

Another gives US corporations permission to export drugs and medical equipment that have been approved in specified developed countries — including Japan, Israel, South Africa and the European Union — but not in the United States.

In terms of overall funding for non-military science and technology, the agreement will leave 1996 spending at about \$32.6 billion, 3 per cent less than the 1995 total of \$34.2 billion. This compares with a cut of 9 per cent in Congress's total non-military budget — so science programmes have fared relatively well. When inflation is taken into account, science and technology spending has been cut by 6 per cent and all spending by 12 per cent.

The budget deal was particularly welcome at the National Science Foundation (NSF), one of several agencies that have been operating without an agreed budget for seven months. According to NSF officials, the agency is \$75 million better off than it would have been if — as was widely feared — no deal had been reached and it had remained dependent on temporary spending measures until the end of the financial year in September.

"I am immensely relieved and pleased that Congress has agreed upon a final budget," says Neal Lane, director of NSF. "We can now put behind us the distractions and confusion of the shutdowns and continuing resolutions," he adds, although he warns that the coming year "may prove to be even more difficult" (see below, left).

The \$221-million budget agreed for the

Republicans outline 1997 priorities

Washington. On the same day that Congress and the White House finally agreed budget terms for 1996, Republicans in the House of Representatives spelled out their science spending agenda for the 1997 financial year, which starts on 1 October.

The Republican plan includes sharp cuts to the National Aeronautics and Space Administration (NASA)'s Mission to Planet Earth, corresponding increases for other space science and significant reductions in climate change and energy supply research.

The new plan was endorsed on 24 April by the House Science Committee, chaired by Robert Walker (Republican, Pennsylvania), which has jurisdiction over most science programmes except for military and biomedical research.

The Omnibus Civilian Science Act proposed by the committee would cut spending on programmes under its jurisdiction from \$20.3 billion to \$19.7 billion. This contrasts sharply with proposals by the Clinton administration to increase such spending to \$20.9 billion.

The act contains the same priorities as the committee put forward last year, although this time proposed cuts of \$270 million in Mission to Planet Earth would be ploughed back into other space science programmes (see box, above right).

The act again tries to eliminate the Advanced Technology Program in the Department of Commerce. But it does allow for a small increase in other programmes at the National Institute of Standards and Technology.

The committee proposes that the National Science Foundation (NSF) be given an increase of 1 per cent for research grants, but that it should be asked to cut administration costs. The NSF would also be told to close one of its seven directorates and — following an amendment by Joe Barton (Republican, Texas) — to change its name to the National Science and Engineering Foundation.

George Brown (Democrat, California), the senior Democrat on the committee, branded the proposals "mean and extreme" and sought, without success, to substitute President Clinton's numbers in their place.

The act is unlikely to pass into law. But if last year's pattern is repeated, its numbers will be accepted by House appropriations committees, who will then have to compromise with the Senate and, finally, the administration. There is already speculation that this will be impossible before November's elections, and that the 1997 budget will have to wait for a new Congress and administration in the new year. **C. M.**

for agencies

ATP, funded through the Department of Commerce's National Institute of Standards and Technology (NIST) is much less than the president wanted. But it does keep the programme alive.

Existing projects will have first call on the money, says Michael Newman, a spokesman for NIST. But the agency hopes, in addition, to be able to launch a general competition for new grants this year.

Most of the final haggling over the 1996 budget concerned not money but the fate of so-called 'environmental riders' attached to budget bills by Congress in an attempt to relax environmental regulation without having to pass separate legislation.

Such riders, which would have curtailed the protection of wetlands by the Environmental Protection Agency, and extended the 'salvage forestry' of old trees, were removed altogether by Clinton.

Others, including the extension of a ban on the listing of new species under the Endangered Species Act, stay in the bill. But Clinton secured the rights for himself, as president, to waive them.

Nevertheless, some riders did take effect, including one that will stop the National Park Service from providing technical assistance to the International Convention on Biodiversity, and another enabling the University of Arizona to override environmental objections and proceed with construction of its proposed Large Binocular Telescope at Mount Graham, Arizona (see below). Colin Macilwain

NASA head gets lean and mean

Washington. The head of the US space agency told his congressional overseers last week that he does not want any more money for space science next year — even if Congress offers it. Reacting to a proposal from the House of Representatives Science Committee for an increase in funding for space science while cutting back on Earth science, Daniel Goldin, administrator of the National Aeronautics and Space Administration (NASA), spoke out forcefully against the plan in testimony before a House appropriations subcommittee the following day.

"I am not supportive of [additional funding for space science], regardless of what happens to Mission to Planet Earth," Goldin told the subcommittee, chaired by Jerry Lewis (Republican, California). He added that he was "vehemently opposed" to increasing money for space science in 1997, as he wants to continue teaching scientists a lesson about fiscal restraint.

Goldin said he wanted the "stress of the budget" to be felt by the space science community. Increasing the budget would send a message that scientists could return to the days of big, expensive missions and would put a chill on his philosophy of "better, cheaper, faster".

Republican staff members on the House Science Committee expressed surprise at Goldin's remarks. They said

later that they did not intend to bring back large, expensive missions, but rather to fund the kinds of innovative technology and science programmes that Goldin has championed.

These include the Discovery planetary missions (which would receive \$20 million more than NASA requested), the New Millennium technology development programme (\$18.5 million more), Explorer projects (\$25 million more) and missions to Mars (\$30 million more).

Democratic opponents of the Republican plan say it is motivated less by a desire to boost space science than to eviscerate NASA's Mission to Planet Earth, particularly the Earth Observing System (EOS) of orbiting remote-sensing spacecraft. Robert Walker (Republican, Pennsylvania), chair of the Science Committee, has always contended that his opposition to EOS is not ideological, but based on his concern that the space agency can not afford the programme as at present conceived.

The committee's proposed 1997 budget would cut \$374 million from a \$1.4-billion request for the Mission to Planet Earth, and kill the Chem-1 mission altogether. The proposal also halves the \$260 million budget for the EOS data system. Goldin spoke angrily against the proposed cuts to EOS in his testimony, saying they would be "devastating to the programme".

Tony Reichhardt

Mount Graham telescope gets green light from Washington

Washington & Munich. The long-delayed US-Italian project to build the large binocular telescope (LBT) on Mount Graham in Arizona might finally proceed, after President Bill Clinton signed a budget bill last week (see above, left) containing a provision specifically inserted to enable the project to overcome environmental objections.

A consortium of German astronomers is also poised to join the project, allowing it to be built to full specification. Officials of the University of Arizona (UoA) will this week ask a district court judge, Alfredo C. Marquez, to lift an injunction blocking construction of the telescope, in the light of the change in the law.

Construction came to a halt in 1994, after Marquez had ruled that a late site change placed the project in breach of the Endangered Species Act (see *Nature* 370, 407; 1994). Environmentalists have long tried to block the building of the LBT, first because of its alleged threat to a local species of red squirrel, and more recently on the grounds that the mountain is sacred to the San Carlos Apaches, a local Indian tribe.

Peter Strittmatter, director of the Steward Observatory at the university, declines to speculate how long it will be before construction work actually begins. Roger Angel, head of the mirror laboratory at UoA, which will build the mirrors for the telescope, says: "I hope we are talking days or weeks, but that is only speculation." Environmentalists intend to continue to fight the project, but have not announced their strategy.

There may be more surprises in store. Early this week, the Mount Graham forest was engulfed in a forest fire during the driest Arizona weather for many years. A spokeswoman for the Forest Service said that the threat to the observatory was "diminishing" as the wind direction changed.

Franco Paccini, director of the Osservatorio Astrofisica in Florence, says that Congress's decision to put an end to the legal harassment of the construction plans is "great news". But he says that the LBT schedule has been so interrupted that a delay of a year might be expected.

Initially planned as a joint US-Italian venture, the project had not been able to

secure the level of financing it had hoped for. Italy will contribute 25 per cent of the costs. The state of Arizona will pay 25 per cent, and other confirmed sources will increase the total US contribution to 42 per cent. The two countries had decided to go ahead with building the telescope with only a single mirror, hoping to find a partner to help complete the second mirror.

Driven by an enthusiastic astronomy community, Germany has now taken on that role. Scientists from various research institutions, chiefly the Max Planck Institute for Astronomy (MPIA) in Heidelberg and the Astronomy Institute Potsdam (AIP), have obtained promises of cash from various public sources.

According to Steve Beckwith, a director of the MPIA, there is now enough money committed to pay for the second mirror, and the partnership will suggest an initial level of participation of 15 per cent. But the German astronomy community would also like to provide instrumentation, and is aiming at a final contribution of 25 per cent.

Colin Macilwain & Alison Abbott

Europe chooses cosmic radiation project

Paris. Cobras/Samba, a new satellite designed to measure cosmic background radiation over the whole sky, has been given top ranking over four rival projects for the next 'medium-sized' space mission of the European Space Agency (ESA), which is scheduled for launch around 2004.

The project was selected at a meeting in Paris last week of ESA's space science advisory committee (SSAC). It will be the third medium-sized mission within ESA's Horizon 2000 science programme, which began in 1985 and is scheduled to run until 2015.

Such missions, which are budgeted at around ECU345 million (US\$424 million) each, complement the larger cornerstone

missions of about ECU625 million that form the backbone of the programme.

The Cobras/Samba mission would cost about ECU340 million, and would have 10 times greater sensitivity and 50 times the angular resolution of the Cosmic Background Explorer (COBE). It was COBE which, in 1992, discovered temperature irregularities in the cosmic background radiation field caused by primordial perturbations occurring within 10^{-35} seconds of the Big Bang.

The new satellite would build on these results. It would aim to establish whether the large-scale uniformity of the Universe resulted from an early period of expansion

known as inflation, and it would look for primordial irregularities that gave rise to galaxies and other structures in the Universe.

As well as providing information on the physics of the early Universe, such measurements can also be used to estimate many astrophysical parameters — such as the Hubble constant — to within a few per cent.

The choice of Cobras/Samba came at the end of a two-day meeting in Paris, at which details of the five projects were presented to about 250 scientists. The advisory committee's recommendations will now go to ESA's science programme committee, which meets in early June. The latter panel, which makes the final decision on the choice of the mission, has never overturned a recommendation by the SSAC.

But any victory celebrations will be dampened by the knowledge that the winning mission must find ways to reduce its costs by 10 per cent, which will mean delaying the launch from the scheduled date of 2003 by at least one year.

ESA decided earlier this year that this mission would be used as a pilot test to assess what savings could be made in mission costs, after the agency had been forced by member states to achieve an annual reduction of around 3 per cent in ESA's science budget (see *Nature* 379, 476; 1996).

Cobras/Samba's suitability for this pilot test seemed to have acted in its favour in the competition with its main rival, Intermarsnet, a project that included three landers built by the US National Aeronautics and Space Administration and a data-relay satellite built by ESA.

According to an ESA official, Cobras/Samba scored points for the fact that its proposed costs are already well within budget, while its design lends itself to a relatively straightforward pilot phase.

In contrast, Intermarsnet was the most expensive of all the five finalists. ESA was also concerned that it would be unable to meet NASA's scheduled 2003 launch date for the mission, given the delays that would be needed to carry out the savings pilot test.

Intermarsnet itself prevailed over MORO — a lunar orbiting observatory aimed at studying the Moon's surface and interior — as the choice of planetary scientists. In the astronomy field, Cobras/Samba was given preference over STARS, a mission to measure the internal structure of a wide range of stars by fine analysis of fluctuations in their brightness.

The remaining candidate, STEP, was aimed at testing the equivalence principle and carrying out a broad range of experiments in fundamental physics. But the proposal suffered from uncertainty about US and French plans for a similar satellite, and the fundamental physics group decided not to push its candidacy. ►

Pentagon delay on asteroid mission

Washington. Funding for a US military project to send an advanced spacecraft to rendezvous with several asteroids is being held up within the Pentagon, even though Congress agreed to provide \$20 million for the mission last year and the Department of Defense gave it the official go-ahead. Work on Clementine 2 has effectively stopped while project managers and their political allies try to persuade the Air Force and the Pentagon comptroller to release the money.

Sources in Congress say they do not know why the funds are being withheld, but some suspect that the Defense Department is trying to 'reprogramme' money to support operations in Bosnia. Members of the Clementine team say that funding problems of this type are not uncommon with advanced military technology projects.

Clementine 2 is scheduled to be launched in May 1998. The plan calls for a spacecraft to fly past three or more near-Earth asteroids, and to release small probes to crash into their surfaces, so that sensors on the main spacecraft can observe the impacts.

Like Clementine 1, which mapped the Moon in 1994, Clementine 2 is intended chiefly to test advanced spacecraft technologies developed for strategic defence and other purposes. One important objective is to prove the ability to track and rendezvous with fast-moving targets in space. The mission also supports a military programme for 'planetary defence' — protecting the planet against possible extraterrestrial threats.

The Air Force's Phillips Laboratory in New Mexico has lead responsibility for Clementine 2, with the Naval Research Laboratory in Washington DC responsible for integrating the spacecraft, and the Lawrence Livermore National Laboratory in California developing sensors.

Project engineers completed a 'concept design review' in March. But the hold-up in funding has prevented further progress. A 'preliminary design review' is due in July

and a 'critical design review' in November; that schedule is now in jeopardy and project engineers have identified several back-up options in case the launch is delayed. A slip of several months, to September 1998, would in fact offer some advantages, including a shorter mission duration (9.5 rather than 13 months) and eliminating the need to use the Earth's gravity to give the spacecraft a boost to reach its asteroid targets.

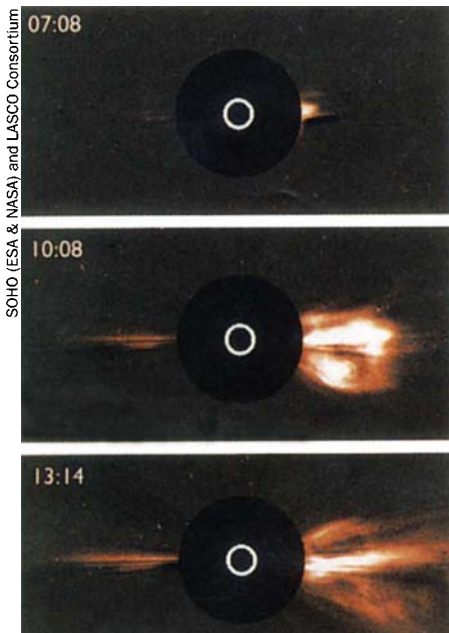
But the September 1998 date would leave no room for slips in the schedule, and the spacecraft would not have as good a viewing angle once it reached the asteroids. Additional delays would also make it more difficult to keep the project within its budget of approximately \$100 million.

The funding hiatus could also slow down the work of an informal Clementine 2 science advisory team headed by Eugene Shoemaker of the Lowell Observatory in Arizona, which, among other tasks, has to continue refining orbits for the target asteroids. The team also needs to begin selecting the science instruments that will fly on the spacecraft.

Shoemaker still hopes that the National Aeronautics and Space Administration (NASA) will participate in the project, as it did in Clementine 1, by sponsoring a formal scientific advisory team. But NASA officials say this is unlikely, given current funding constraints. NASA also has its own low-cost asteroid mission scheduled for 1998 and NEAR (Near-Earth Asteroid Rendezvous), due to meet the large asteroid Eros in 1999.

No one is certain how and when the funding issue will be resolved. James Muncy, a congressional staff member who works on space issues for Dana Rohrabacher (Republican, California), says that several members of both the House and the Senate are working on the problem. He adds that it is simply a matter of continuing to put pressure on Pentagon officials until they release the money.

Tony Reichhardt



Images released today from the joint ESA-NASA SOHO project show a coronal mass ejection: billions of tonnes of gas burst out of the Sun at 550 km per second.

► Cobras/Samba and Intermarsnet entered the final selection round neck-and-neck according to scientists at the meeting. In particular, the meeting had to digest the significance of a recent US announcement that it intended to launch a similar mission to Cobras/Samba, called MAP.

But the Paris meeting concluded that MAP was "definitely an inferior mission", according to a senior official of the space science programme, who says that the performance of the US satellite had been "overstated".

In order to prevent the portfolio of ESA's missions early in the next decade being dominated by astronomy, the SSAC recommended that the next cornerstone mission be reserved for a planned mission to Mercury that would include an orbiting observatory and a landing craft (see *Nature* 371, 643; 1994). It also recommended that the next medium-sized mission be reserved for a project on Solar System research. "The success of Cobras/Sampas has been paid at a certain cost by the astronomy community," says one ESA official.

Declan Butler

German research loses immunity to cuts

Munich. German research minister Jürgen Rüttgers has not maintained the financial advantage he gained for his ministry last year when he secured an above-average budget rise of 2.8 per cent for 1996 over the 1995 level. The research ministry will not be exempt from a new round of budget cuts being imposed on all government spending as part of an emergency budget, introduced to combat the country's budget deficit resulting largely from the continued costs of reunification.

The federal government agreed last

Geneticist leaves Cambridge for industrial research post

London. One of Britain's leading geneticists is leaving a senior academic post to head the genetics research activities of SmithKline Beecham (SB), the world's tenth largest pharmaceutical company, which says that it is planning to base all future pharmaceutical and diagnostic products on knowledge of the human genome.

Peter Goodfellow, Arthur Balfour professor of genetics at the University of Cambridge, and well-known for a wide variety of contributions to genetics research — ranging from somatic cell genetics to sex determination — has announced his resignation, and will join SB at its main British research site at Harlow, outside London, at the beginning of July.

Goodfellow's arrival is being seen by the company not only as laying an important cornerstone for its genomics strategy, but also as evidence of a growing overlap between academic and industrial researchers, the distance between which has traditionally been greater in the United Kingdom than in countries such as the United States.

One aspect of this overlap, according to some, is that the cost of advanced research facilities in fields such as DNA sequencing is now so high that the private sector can offer more attractive basic research facilities than the academic sector. Goodfellow, for example, will have access to the privately held database of genomic sequences built up by Human Genome Sciences and the Institute for Genomic Research in Rockville, Maryland, in which SB has a major share.

"We are delighted to be bringing some of Goodfellow's stature into the organization," says George Poste, director of research and development for the pharmaceutical company, formed in 1989 out of a merger between the US company SmithKline French and Britain's Beecham.

Poste has been largely responsible for a strategy that aims to make knowledge of the genome the starting point for all new diag-

nostic and therapeutic products in SB by the end of the decade. Sequencing the human genome, he argues, is only the first step; equally difficult is learning how to exploit the knowledge that this produces. "The whole world of science has to spend a lot of time trying to assess the functional role of genes," he says. "Goodfellow's appointment is an expression of that reality."

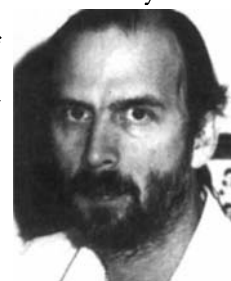
Genomic knowledge, he says, will not be used only as the basis of new diagnostic tests and potential therapies. Another potential application is in the testing of new drugs before approval by regulatory authorities, attempting to understand the 'pharmacogenetic' basis of drug response at the molecular level — in other words why different people respond to drugs in different ways.

Goodfellow already has industrial experience. He has, for example, been a consultant for Amersham International and a founding member of the scientific advisory board of Sequana Therapeutics in California, which recently announced the discovery of the 'tubby' gene responsible for one type of obesity in mice (see *Nature* 380, 534-538; 1996).

He expresses both regret and relief at leaving the academic environment. "I've had a wonderful time at Cambridge," he says. "But there have been constraints. It has been a strain to maintain the infrastructure needed to carry out internationally competitive research and the financial restraints have been getting worse."

He will also be giving up the sex determination work that he has been involved in for the past ten years. But at the same time, he says that his new UK-based post, in which as a senior vice-president for research and development he will be one of the managers responsible for SB's worldwide research, will involve "taking the sort of science that we do in universities into an industrial context".

Meanwhile, Poste says that he also hopes that Goodfellow's arrival at the company will help to boost the morale of its scientists in Britain which, he admits, was badly shaken by his decision last year to close down part of its microbiological activity, and which is further undermined by talk of future job losses among research teams considered insufficiently productive. "This sends a signal that we are committed to the long-haul in the United Kingdom," he says.



Goodfellow: regret and relief to be moving on.

David Dickson

British politicians press case for Human Genetics Commission

London. A British parliamentary committee has reiterated a suggestion first made last year that the government should set up a broad-ranging Human Genetics Commission to deal with the inter-relationship between many different aspects of genetic science. It says that it is confident that this suggestion — initially rebuffed by Whitehall — will now be accepted.

The proposal was made by the House of Commons select committee on science and technology, following a lengthy inquiry into the field (see *Nature* 376, 202; 1995). In its response to the committee's conclusions last year, the government rejected the idea of a new body as unnecessary, in the light of existing specialist advisory committees. But in a highly unusual move, the committee organized a follow-up series of hearings to restate its case.

During these hearings, government ministers indicated that they were prepared to reconsider their earlier reaction (see *Nature* 380, 6; 1996), and in a report on the hearings published last week, the all-party committee says it now expects some form of body to be set up by the government. The main aim of the body would be "to foster public confidence and understanding" of genetics. But it adds that "in practical terms the commission's function would be to advise the government on the broad issues raised by genetic science." □

Genetic data ban for US insurers

Washington. The US Senate last week unanimously passed a health reform bill that explicitly bars insurers from using 'genetic information' to deny coverage to applicants. The bill's House of Representatives counterpart, passed on 28 March, also includes such information among factors that insurers may not use to deny coverage, making it highly likely that both bodies will agree to this measure in any joint legislation. The main goal of the Senate bill, whose authors are Nancy Kassebaum (Republican, Kansas) and Edward Kennedy (Democrat, Massachusetts), is to prevent people from losing health insurance when they change or leave jobs.

Kassebaum initially resisted modifying the bill to include 'genetic information' among the factors insurers may not use to turn potential clients away (see *Nature* 380, 91; 1996). But at the urging of scientists and patient and other advocacy groups, she inserted the two words in the bill before it came to a vote on 23 April. □

Solar telescope performs to plan

Paris. The Franco-Italian telescope Thémis — Télescope Héliographique pour l'Etude du Magnétisme et des Instabilités Solaires — which will be officially inaugurated in June, has reached its planned performance levels, according to results of the first spectra obtained using the telescope. Two-thirds of its cost of FF87 million (US\$11 million) have been met by France's Centre National de la Recherche Scientifique, and the rest from its Italian counterpart, the Consiglio Nazionale delle Ricerche. Providing information on the solar wind and solar turbulence, the telescope will complement studies of the solar corona by the SOHO satellite launched last December by the European Space Agency and NASA (see page 9). □

University cuts 'harming health'

London. Budget cuts to universities that train Britain's doctors will have an adverse effect on the National Health Service, according to the UK Committee of Vice Chancellors and Principals (CVCP), the heads of the universities. At a press conference in London last week, the CVCP said that the 31 per cent reduction in universities' capital income announced in the last budget would damage medical research and potentially imperil the lives and health of patients.

The budget cuts have already delayed a new institute of health sciences in Oxford, according to the CVCP. Similarly, the Institute of

Psychiatry at the University of London is unable to fund a new building to house the Social, Genetic and Development Psychiatry Research Centre. The government has increased the annual intake of medical students by 500 to 5,000 in order to meet a predicted shortfall in doctors. But Sir Michael Thompson, chair of the CVCP's medical committee, said there was no sign of the extra money needed to teach these students. □

'Racist' psychologist in book row

London. A psychology lecturer at Edinburgh University is demanding \$75,000 after publisher John Wiley & Sons Ltd withdrew his book from publication. The move by Wiley & Sons follows reports in the British press that Christopher Brand, author of the book concerned — *The g factor: general intelligence and its implications* — described himself as a "scientific racist". Brand says Wiley has offered to release to him unjacketed copies and reassign copyright. □

East-West projects earmarked

Moscow. A joint committee set up by the European-funded International Association for Promotion of Co-operation with Scientists from the New Independent States of the former Soviet Union (INTAS) and the Russian Foundation for Basic Research (RFBR), meeting in Moscow, has identified 145 projects to receive overall financial support of 6.6 million ECU (US\$ 8.1million).

The average funding of each project will be about 50,000 ECU over two years. A total of 1,353 proposals received up to the middle of last December were subject to peer review in a first round of evaluation. Each project involves at least four scientific teams, two from Russian scientific institutions and two from different INTAS Member States. At least 80 per cent of the funds are allocated to the Russian research teams. The results of the joint evaluation will now be submitted to the respective decision-making bodies of both INTAS and the RFBR. □

Canada plans single blood agency

Ottawa. Apparently yielding to public concern stirred up by controversy over how decisions over HIV-tainted blood were made in the mid-1980s, federal and provincial health ministers have agreed to reorganize Canada's blood system under a single new agency. David Dingwall, the federal health minister, says there is a consensus among the ministers that an agency is needed that is independent from political considerations and with clear overall authority. A decision on the agency's final shape will be made by September. □

Live surgery on the Internet

London. A surgical operation has been broadcast live on the Internet for the first time. Transmitted to an estimated 60,000 users through the Health Online Service, the operation at the German Cancer Research Centre in Heidelberg involved two techniques to replace cruciate ligaments. Images were displayed alongside commentary from one of the surgeons, and viewers were able to send questions via e-mail. □

Joint research in South China Sea

Melbourne. The Philippines and Vietnam are to send their first-ever joint marine research team to the disputed Spratly Islands in the South China Sea. The 15 Filipino and 12 Vietnamese scientists say that they will compile an inventory of marine organisms during the 15-day trip aboard a Philippine government boat.

The territorial waters of the Spratlys are believed to include large deposits of oil and gas. The six countries that claim sovereignty over the islands — Taiwan, Vietnam, Brunei, Malaysia and the Philippines — have agreed to conduct joint research as a means of reducing military tension, and a first step towards greater multilateral cooperation. □

Whose genes are they anyway?

The increasing interest of pharmaceutical companies in the results of genetic studies of human population groups is already raising intense debate over the 'ownership' of genetic information. Few expect to find easy solutions.

London. Imagine the scenario. A group of individuals — they could be South American Indians participating in an anthropology project, inhabitants of villages near a nuclear power plant, or patients suffering from a debilitating, inherited disease — agrees to provide researchers with blood samples.

At the time, the terms of the agreement are relatively informal. The Indians accept some medical equipment and help in building a local school, the villagers a promise that they would be contacted individually if the samples reveal evidence of medical problems, the patients a commitment that their blood will contribute to a treatment for their disease.

Three years later, the rights to genetic data obtained with the help of the samples in question are sold by the researchers to a large pharmaceutical company for, say, \$20 million (the rights to one gene associated with obesity were sold last year for \$70 million).

What could happen next? The Indians might take the researchers (or the company) to an international court to claim a share of the royalties; the villagers sue for the misappropriation of their data; the patients demand royalty-free access to any resulting diagnostic techniques. The possibilities are wide, and already taxing the minds of those who accept that each scenario is far from implausible.

"The issue is a nightmare", says one geneticist who works for a large pharmaceutical company. "This whole area of people's blood being used to find genes which are then commercialized could become a watershed in the relationship between pharmaceutical companies and health care in general."

Some argue that the issues involved are not particularly new. They point out that patients have taken part for many years in clinical trials of pharmaceutical products with no expectation of any financial return — and that the contribution a DNA sample from any one individual to a significant discovery is likely to be relatively small.

"If a journalist writes an article on a family, and then wins a Pulitzer Prize, does he give the family a percentage his winnings?", asks Arnold Slutsky of the Samuel Lunenfeld Institute of the Mount Sinai Hospital, Toronto, who has been involved in using samples from patient groups in various locations — including the island of Tristan da Cunha in the South Atlantic — in a search for genes involved in asthma. "It is not really that different."

Others point to recent interpretations of

patent law on both sides of the Atlantic to argue that, whatever the ethical issues, individuals have no further claim over tissue taken from them with their consent.

But three factors appear to have made the growing controversy over genetic information qualitatively different from past debates. The first is the symbolic nature of the gift of blood in both developed and 'primitive' societies. Many fear that the gift relationship, including donations by individuals of samples for use in medical research, may be corrupted by the intrusion of financial interests.

A second factor is a broadening aware-



The family way: Indigenous populations, such as these in Sri Lanka, offer rich prospects for geneticists.

ness of the financial value that might attach to some genetic sequences, of which the obesity gene is an extreme but by no means unique example. Indeed, the activities of researchers looking for new genes is now frequently referred to as 'gene-prospecting', even by those who support such activities.

Third, many field researchers say they are now encountering communities who feel they have been exploited by large agricultural companies which are selling back to them seeds initially collected in seed banks in the 1960s and 1970s. Such communities frequently use this experience to justify their reluctance to participate in contemporary efforts to build up 'human gene banks'.

All three factors have, for example, come into play in attacks on the Human Genome Diversity Project (HGDP), a plan to collect DNA samples systematically from about 500 linguistically distinct groups throughout the world in order to analyse the dynamics of human genetic variation.

Among the most vocal critics of the HGDP have been environmentalists and groups in developing countries, veterans of earlier campaigns to challenge the increasing control over the world's food crops exercised by a relatively small number of large — and mostly Western-based — agricultural and seed companies.

One such group is the Rural Advance-

ment Foundation International (RAFI), based in Ottawa, Canada, which is leading a global campaign designed to draw attention to the potential implications of the HGDP. The group has publicly called for the project to be stopped, at least until strict controls have been agreed on the way in which any information emerging from it is used.

It was RAFI researchers, for example, who first came across — and widely publicized — the fact that the US Patent and Trademark Office last year had granted a patent to the US Department of Health and Human Services on a human t-lymphotropic virus (HTLV-I) derived from the Hagahai people of Madang Province in Papua New Guinea.

RAFI's criticisms are sharply rejected by many of those involved in developing the diversity project itself. Luigi Luca Cavalli-Sforza, for example, the population geneticist largely responsible for conceiving the HGDP, claims that its goals have been misinterpreted and exaggerated (see page 14).

The HGDP has recently taken pains to distance itself from possible disputes about patent rights by pointing out that it is not accepting support from commercial sponsors. HGDP participants such as Henry Greely, professor of law at Stanford University in California, have been watching carefully negotiations involving patent rights, such as that which led to an agreement last month between Carol Jenkins, the anthropologist involved, and the government of Papua New Guinea (see *Nature* 380, 374; 1996).

Jenkins, who as 'co-inventor' on the patent was entitled to 50 per cent of the royalties from any resulting commercial products, had already agreed to give her entire share to the Hagahai people, and says may do this through formally through assigning her royalties to a neutral body that would manage any proceeds to the benefit the indigenous group. Greely says that this represents a possible approach to the HGDP's pledge to return a fair share of any commercial returns to participants in research.

Yet RAFI's campaigning efforts have ►

On other pages

- Population groups hold the key 12
- India wonders what to do 13
- The diverse human genome 14

For reprints of this briefing, contact Ross Sturley, *Nature*, 4 Crinan Street, London N1 9XW, UK

Population groups can hold critical clues

London. As geneticists try to unravel the basis of ever more complex disorders, the search for informative population groups—whether defined ethnically, geographically or by common medical history—is assuming increasing importance.

Several thousand human diseases exhibit classic mendelian (that is, recessive or dominant) inheritance, and are presumably caused by defects in single genes. But many other single-gene diseases do not follow strict mendelian patterns of inheritance. In these cases, environmental factors influence who gets sick, and the terms 'genetic predisposition' and 'susceptibility genes' are often used. Not all individuals with a particular defective gene will thus develop disease, although those who do are overwhelmingly likely to have inherited the defective gene.

Other disorders are caused by the interaction between several genes; in such 'complex disorders', the disease can be caused by a collection of genetic alterations, none of which is in itself harmful, and any one of which only slightly increases an individual's risk of disease. Here in particular, the disease does not follow simple mendelian inheritance patterns; there is no simple correlation between genotype and phenotype, and identifying the genes involved is a difficult task.

Unfortunately, most common diseases such as diabetes and asthma, as well as major killers such as cancer and heart disease, are complex disorders. But the hope that unravelling their genetic basis will lead to improved treatment or ultimately a cure has made the hunt for the genes involved a priority for researchers in academic institutions and the pharmaceutical industry.

Whatever the condition, the gene must be singled out from 80,000 others on the human genome, a feat usually performed by identifying where on the genome the gene lies. The process requires both raw material, in the form of DNA samples from related groups of individuals in whom the disease is common, and genetic markers—highly variable short sequences of DNA at unique, known locations.

Such markers can be used to map disease-causing genes in many ways. All rely on the idea that if a marker lies close to the gene in question, one form of the marker is likely to be inherited along with the gene in an affected family or inbred population.

Once two markers flanking the gene have been found, they can be used as 'entry-points' for cloning DNA from the region, which in turn can be used as a source of candidate disease genes. Comparing the DNA sequence of such genes in affected and unaffected relatives leads eventually to identifying the gene whose

defects are associated with the disease.

In principle, it should be easy to tell whether a gene is defective only in affected individuals. In practice, the human genome contains many polymorphisms (different forms of the same gene), and the sequence of a gene will often vary slightly from person to person unless the individuals in question share a recent common ancestor.

Most of these differences do not affect gene function. But their frequency can make it hard to identify a disease gene on the basis of sequence alterations alone.

Analysis of data from the more complex mapping techniques generally requires sophisticated computer software, and even then, very large areas of DNA must often be sequenced to identify the gene in question.

Vital samples: blood supplies key data.

But the number of known markers, the quality of the software and the speed of sequencing have all improved substantially over the past five years. Most geneticists now consider the availability of suitable DNA samples to be the limiting factor in the study of genetic diseases.

In the past, geneticists have concentrated on single-gene diseases. Using extensive pedigrees, they have identified how the disease is inherited and compared the genotype of affected and unaffected individuals. About 500 human genes have been mapped in this way, more than 10 per cent of which have been successfully cloned.

The study of complex disorders is more difficult, as unaffected individuals may still carry the predisposing genetic alteration. Mapping studies must therefore concentrate on affected relatives, searching for genomic regions that they share more often than random inheritance would predict.

In affected families, the method most commonly used is the analysis of affected sibling pairs, known as 'sib pairs'; a recent search for genes affecting susceptibility to diabetes involved over 500 such pairs. Sib-pair analysis is now being carried out for most complex diseases.

But finding a sufficient number of sib pairs for a particular disease can be difficult, and the chromosomal region defined is seldom small enough to make identification of the gene involved easy.

Researchers are therefore turning increasingly to genetically isolated populations. Such populations often have only a small number of founding members, and a certain degree of inbreeding. Individuals therefore share significant parts of their genomes (although variation increases with time), and genetic alterations predisposing to disease tend to be inherited together with their neighbouring DNA.

As with sib-pair analysis, scanning the genomes of affected individuals for conserved chromosomal regions can be an efficient way of mapping genes of interest, sometimes down to a region small enough to allow cloning.

Nor is this the only advantage offered by genetically isolated populations. For example, variations in any one component of a complex system may be more obvious if the other (genetic and environmental) components are constant among most individuals in a group, making susceptibility genes easier to spot. A lower level of 'background' polymorphism can also help to single out such genes from among several candidates.

The Finns are a prominent example. The founders of the population arrived in Finland only about 2,000 years ago (75 generations), and the population went through a drastic decline as recently as the late seventeenth century, implying a further reduction in overall polymorphism.

Until recently, the Finns mixed little with other groups, thus limiting their genetic diversity. Furthermore, Finland has exceptional family records dating back to the seventeenth century.

Such favourable conditions have already led to the identification of genetic defects underlying several single-gene diseases in the Finnish population, and there is general optimism that this population will also be useful in the study of complex diseases such as schizophrenia and diabetes.

The advantages of such populations have encouraged scientists and pharmaceutical companies to search systematically for new remote populations with high incidences of particular disorders. One example that promises to help find asthma genes is the highly inbred population of Tristan da Cunha, a small island in the South Atlantic, one third of whose roughly 300 inhabitants suffer from the disease.

Not every disease will have a 'perfect' island. But countries such as India and China, with remote regions where population groups have little geographical mobility and a high degree of inbreeding, may well prove to be genetic goldmines. The ethical and economic problems posed by exploiting such resources will therefore be an inescapable feature of genetic research for some time to come. **Barbara Cohen**

Luke Holland

► struck a sympathetic chord among many indigenous groups, for whom the issue of DNA sampling by Western scientists, particularly against the background of growing interest in the analysis of such samples by pharmaceutical companies, brings back sensitive memories of the colonialist past.

Last year, for example, a group of countries in the South Pacific, alarmed by reports such as that from Papua New Guinea, drew up proposals for what it is calling a 'patent free zone'. Less confrontationally, but equally significantly, the government of India is developing legislation to require foreign researchers using genetic material taken from India to ensure that any resulting technical advances — as well as a share in the profits from the eventual exploitation of this material — are transferred back to India (see right).

Such a practice, endorsing the concept that those who provide DNA samples to researchers are entitled to a share in any royalties from their use, represents a major shift in thinking. Until now, the 'exchange' of blood samples has been made on more pragmatic — and often less formal — grounds, involving for example either token 'gifts' in return, or practical help such as medical supplies and treatment.

Yet some claim that the very process of providing a financial reward (in terms of royalty payments) could threaten the cultural traditions of indigenous groups. "It would be ironic if through this approach to royalties we end up contributing to the process of 'commodification' of individuals", says Bartha Knoppers, a professor of law at the University of Montreal in Canada.

The debate is not confined to developing countries. Even in developed countries, where there has traditionally been little difficulty in finding volunteers for medical research projects — particularly where such research is seen as contributing to the development of medical services — genetics researchers report a growing wariness about the implications of a growing commercial dimension.

Researchers in Finland, for example, whose stable and well-documented social structure has provided a valuable resource for genetic studies, point out that the willingness of families to participate is based partly on trust in the state-run health-care system — as well as knowledge of its cost — and that this trust could be lost if the family data was to be 'sold off' to some drug company.

Few see any quick resolution to the sensitive and complex issues involved. One potentially significant move took place last month when the Human Genome Organization (HUGO) endorsed a set of principles drawn up by its ethics committee on the conduct of genetic research (see *Nature* 380, 279; 1996).

In a key passage, the statement states that 'undue inducement' through compensation for individual participants, families and populations should be prohibited. But it also specifies that such a prohibition should not ►

Gene-hunters home in on India

New Delhi. In the past, Madhav Gadgil, professor of ecological studies at the Indian Institute of Science in Bangalore, had no hesitation about sending blood samples from various indigenous groups — including high-caste brahmins, low-caste *hijras* and a hunter-gatherer tribe called *kadar* — to population geneticists at Stanford University in California.

But Gadgil says the situation has now changed, in particular following the country's decision to sign the General Agreement on Tariffs and Trade (GATT), and other events that have raised questions in India about the extent to which human genetic material should be patented. "I have now declined to send any further samples to the United States until our government's position is clear," he says.

It is an issue about which feelings run high, fanned by the growing interest of the world's pharmaceutical companies in genetic information. Kiren Kucheria, a geneticist at the All India Institute of Medical Sciences (AIIMS) in New Delhi, says she received an offer — which she refused — of US\$20,000 for blood samples from two patients suspected of having a 'dislocated' gene (a gene that has switched places with another on a different chromosome) for müllerian ductagenesis, a disorder characterized by the absence of a uterus.

Perhaps unsurprisingly, such reports have fuelled the efforts of groups keen to build on the popular political appeal of earlier campaigns about the exploitation of India's natural environment — and, more recently, of its plant genetic resources — by foreign multinationals.

"A lot of blood and DNA samples, from tribal people and clinically identified cases, are leaving India without permission," says Suman Sahai, a geneticist who leads the Gene Campaign movement. "They will eventually become raw material for the pharmaceutical industry without our people getting any benefit."

Indeed, some scientists claim that foreign researchers are scrambling to collect human genetic material before the government tightens its laws. K. Suresh Singh, until recently chief of the Anthropological Survey of India, says that he has recently received so many proposals from universities in the United States and elsewhere for field studies among tribal people that he has had to reject most of them.

Many scientists accept that foreign involvement does have some important advantages. Kucheria says that collaboration with the Institut Pasteur in Paris helped to build her laboratory and to train students at the AIIMS. Others point out that any treatment arising from such stud-

ies will eventually be available — in principle — to people suffering from diseases.

Collaboration projects with foreign research groups involving the export of human blood and DNA samples at present require the approval of a small committee within the Indian Council of Medical Research (ICMR). But many collaborative projects appear to ignore — or be unaware of — this requirement (see *Nature* 371, 381–382; 1996).

The Department of Biotechnology (DBT) took the first step towards a more thorough approval procedure in March when it received approval from the health ministry to expand the existing committee in the ICMR to include representatives from other scientific agencies involved in research in human genetic material.

Applications for foreign collaborations and export of genetic material will have to be approved by the committee, and violators will be liable for penalties under legislation due to be introduced in the next session of parliament. According to Subash Chandra Dey, a senior official at the Ministry of Environment, the law would make export of all genetic material — including human material — conditional on foreign collaborators agreeing to transfer the technology and to pay a share of any profits that may accrue from any discovery resulting from use of the material.

To qualify for patent protection, details of every sample leaving the country will have to be recorded. P. K. Ghosh, a DBT adviser, says the aim is not to stifle collaboration. "What we want is maximum advantage financially for our scientists, the government and the people."

It is considered too early to work out details about sharing because, says Sharma, "commercial spin-offs are only a distant prospect". Nevertheless, the question of intellectual property rights is high on the agenda, primarily because India is struggling to amend its Patent Act of 1970 to bring it in line with GATT.

Officials maintain that having accepted GATT, India has to fall in line with the rest of the world. But two attempts to do so in the parliament have already failed, and opposition parties have been heavily influenced by pressure from environmental groups and other critics on the specific issue of gene patents.

The critics include some prominent figures. "Patenting human genetic material is totally unethical," says M. S. Swaminathan, former director general of the Indian Council of Agricultural Research. "Commercial reward for investment in human genome research must come from the sale of vaccines and medicines and not from patents." **K. S. Jayaraman**

► cover either various forms of technology transfer or "the possible use of a percentage of any royalties for humanitarian purposes".

John Fleming, director of the Southern Cross Bioethics Institute in Australia, and a member of an international panel set up by Unesco to draw up an international code of bioethics, suggests that there is a need for a body to act as an 'honest broker' to help set up agreements covering topics such as patent rights between researchers and the groups they are studying.

Others want a more political solution. Several Pacific nations, for example, are said to be contemplating asking the UN General Assembly to seek an advisory ruling on the morality of human gene patents from the International Court of Justice.

According to Pat Mooney of RAFI, the

ultimate goal of such a move would be to persuade governments to introduce tighter restrictions on human genetic material in the intellectual property provisions of the General Agreement on Tariffs and Trade (GATT), which are due for review in 1999.

But such moves, too, are likely to be highly controversial. Biotechnology companies are already uncomfortable with provisions in the European Patent Convention that allows moral issues to be used to judge eligibility for patentability. Action at the global level allowing for enhanced protection of human genetic material is likely to meet with even stronger opposition.

In contrast, any such proposals are likely to win enthusiastic support from indigenous groups that see the prospect of participating (as subjects) in the search for new genes not

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Navajo Indians: will genetic analysis help them achieve a better future?

as a welcome chance to boost the world's biotechnology industry, but more as a form of modern-day colonialism. **David Dickson**

Diversity project: Cavalli-Sforza answers his critics

San Francisco. Much of the current controversy about the implications of the genetic screening of population groups has focused — some would argue unfairly so — on the Human Genome Diversity Project (HGDP), a programme of research conceived by Luigi Luca Cavalli-Sforza, the population geneticist and professor emeritus at Stanford University in California.

Cavalli-Sforza, the author of books covering a range of aspects of both biological and cultural evolution, combines genetic, archaeological, linguistic and other cultural data in understanding human evolution. The HGDP, he says, would be a collaborative effort to set up a rational method to study human variation and reconstruct the history of modern ethnic differentiation, as well as providing other scientific and applied benefits.

New techniques of DNA analysis, Cavalli-Sforza points out, provide an unprecedentedly precise way of analysing the history of human populations. Using such techniques, for example, he has been collaborating since 1984 with Judy and Ken Kidd of Yale University on a pilot project to generate cell lines from two dozen populations throughout the world. So far, more than 100 restriction fragment length polymorphisms (RFLPs) and 100 microsatellite markers have been tested on 15 cell lines, representing populations from each of the five continents.

"Population genetics has been going on since 1917. The HGDP is an effort to make it more efficient, specific and rational," he says, adding that the increasing rate at which populations are mixing (and losing) their genetic uniqueness gives the project considerable urgency.

In all, Cavalli-Sforza foresees sampling about 10 per cent of the world's 5,000 linguistically distinct populations. One use of this information, he says, would be to quantify how language and other cultural

characteristics correlate with, and may have influenced, the distribution of genetic profiles throughout the world. He says a careful analysis can do much to counter racist interpretations of such correlations.

He also argues that the data would help to provide a more balanced perspective of humanity's genetic resources than the Human Genome Project, which so far is focusing on the genetic profile of a mosaic of a few groups of individuals primarily of European origin. Finally, it could yield important information about the genetic basis of disease susceptibility and immunity.

In contrast, Cavalli-Sforza says that the HGDP has no intention of emphasizing research on populations in danger of extinction. This concept, he says, emerged mostly from the media's mistaken interpretation of the project's interest in collecting information rapidly from populations undergoing increasing admixture — a process that in some cases enhances their fertility and survival, but contributes to the disappearance of cultures and makes genetic measurements more difficult.

Indeed, he says that researchers had hoped that a few populations might provide some interesting genetic surprises, and that drawing attention to the plight of some small tribes might attract efforts to help them survive. Genetic research on such groups, he adds, might also help to solve specific problems, such as the low fertility of the Andaman Islanders of India.

Cavalli-Sforza is disturbed by accusations that his work contributes to racist stereotypes. In fact, he says, genetic evidence strongly supports the concept that there are no separate races.

Although the HGDP has come under intense public scrutiny (see *Nature* 377, 373; 1995), Cavalli-Sforza maintains that this is not unusual for genetic research programmes. But he still admits to being taken aback by the fervour of some

attacks on the HGDP, claiming that the project has been misrepresented by its critics — for example, those who have dubbed it the vampire project because it is based on the collection of blood samples — and had not been prepared to correct their interpretation.

The response to scientific projects that raise difficult questions should not be to bar them, but to approach them with care, he says. In this vein, he has already asked the United Nations Educational, Scientific and Cultural Organization to review ethical issues related to the project, in particular compensation in case of commercial applications of the research.

At the same time, he says that he has chosen not to reply directly to some of his critics because he feels they are politically and personally motivated, and not primarily interested in solving genuine problems.

On the question of patents on genetic information obtained during the course of the project, Cavalli-Sforza says his personal view is that there should be no patents on DNA at all. But he feels that the potential commercial value of the information now emerging from the Human Genome Project and related activities may make such a position impossible to sustain.

Indeed, with indigenous people gaining in political influence, he suggests that patents could provide a way for them to exploit their own resources. "In the unlikely event there is some gene that becomes commercially valuable, the people who donated it — not the individual, but the group — should somehow share in the advantages," he says. But, he adds: "How to do it is not so easy."

One decision that he believes is essential, however, is that the HGDP should not accept any corporate funding, precisely to avoid any potential conflict of interest.

Sally Lehrman

Moving the British cattle herd

SIR — One of the more extreme suggested solutions to the UK bovine spongiform encephalopathy (BSE) crisis is to move all cattle to new pasture after eradication of the disease. Although there may be little scientific basis for such an exercise, political pressure to take radical steps to restore public confidence in the British beef market continues to grow. Furthermore, with the suggested slaughter of a portion of the national herd, some UK grassland may be ploughed under and converted to arable land. We have attempted to quantify the possible impact on soil carbon (C) content and on nitrate leaching of the more extreme of these scenarios, moving all cattle in the United Kingdom to new pasture.

We estimate from statistics of the total number and age distribution of various classes of cattle in the United Kingdom¹ and the area of grassland required by individuals of each class² that the 11,833,500 cattle require at least 4.41×10^6 ha of grassland. The total UK area of permanent grassland³ is 11.05×10^6 ha and the ratio of grassland more than five years old to that under five¹ is 3.72:1. Assuming that the use of grassland by cattle is distributed evenly between permanent and temporary grass, cattle are estimated to use 3.22×10^6 ha of permanent grassland, 29% of current permanent grassland would need to be ploughed under. This may be a conservative estimate as it assumes that cattle use the same fields every year. Using a mean figure for organic C content of grassland soils of 3.36% (derived from figures for six major soil types⁴) and a bulk density of 0.97 g cm^{-3} (from regression against organic C content⁵), the total organic C content of all UK permanent grassland to 30 cm can be calculated as $1077 \times 10^6 \text{ t C ha}^{-1}$. The 29% of this used by cattle contains $312 \times 10^6 \text{ t C ha}^{-1}$.

Long-term experiments in the United Kingdom show that converting permanent grassland to arable agriculture can result in

a 20.3% loss of soil organic C over 15 years⁶. If a similar loss occurred when grassland was ploughed under, $63.4 \times 10^6 \text{ t C}$ would be lost from the terrestrial C pool over 15 years, equivalent to 6% of current C in grassland, or 0.3% of the total soil C stock of $21\,784 \times 10^6 \text{ t C}$ in Britain⁵.

The loss would not be fully compensated for by the increase in soil C stock under new grassland converted from arable land. Long-term experiments show that organic matter increases much more slowly under reseeded grass than it is lost from ploughed-under grassland⁷. Arable soils contain less organic carbon than do grassland soils, so the predicted increase in organic C of 17% over 15 years⁶ is less than that lost by ploughing-under grassland. Using similar calculations, assuming an average arable soil organic C content of 1.37% and a bulk density of 1.2 g cm^{-3} , the total increase in soil organic C for the 53% of current arable land³ ($6.08 \times 10^6 \text{ ha}$) that would be required for new grassland is $27.5 \times 10^6 \text{ t C}$, equivalent to 0.1% of total soil C stock for Britain.

The net change in total soil C content (decrease from ploughed-under grassland minus increase in reseeded arable land) is calculated as $36 \times 10^6 \text{ t C}$ over 15 years or 2.4 Tg yr^{-1} . This level of carbon loss from soils is equivalent to about 1.5% of the annual UK anthropogenic carbon dioxide-C production⁸. In the light of a commitment to stabilize or reduce carbon emissions to the atmosphere, such a loss of carbon from soil is not negligible and would push UK 1991 carbon dioxide emissions to a level not reached since 1980⁸.

In addition to the implications for soil organic matter and loss of carbon to the atmosphere, the ploughing under of grassland would be likely to result in a significant increase in nitrate leaching. Ploughing-under permanent grassland can lead to increases in nitrate leaching of 4 t N ha^{-1} over 20 years⁹, half of which can occur in the first 5.5 years. This is equivalent to $363 \text{ kg N ha}^{-1} \text{ yr}^{-1}$ during the first 5.5 years and could lead to concentrations in the water draining from fields in the first season⁹ of up to 450 mg N l^{-1} depending on rainfall. This would almost certainly push the nitrate concentration in drinking water above the European Commission limit in some areas. Over the whole of Britain, the ploughing-under of the permanent grassland required to move all cattle could result in the leaching of an extra $4.4 \times 10^6 \text{ t nitrate-N}$ over 20 years.

The adverse environmental impact would also be partially realized if any shift in land-use from grassland to arable crops occurred as a result of selective cattle slaughter. Direct and indirect effects in

other areas of environmental concern (for example, trace gas fluxes from the soil) are also likely.

Our preliminary calculations suggest that there would be significant and long-lasting adverse environmental effects associated with the ploughing under of grassland in response to the current BSE crisis. Our calculations suggest that in the absence of a sound scientific basis, the ploughing under of permanent grassland in order to move the national British cattle herd would be extremely unwise.

Pete Smith

Jo U. Smith

David S. Powlson

Soil Science Department,
IACR-Rothamsted,
Harpenden, Herts AL5 2JQ, UK

Plateau reached

SIR — The letter from Olle Persson of Inforsk about data from the Institute for Scientific Information (ISI) relevant to trends in multinational collaboration (*Nature* **380**, 100; 1996) carried a particularly unfortunate title ("ISI miscount?").

Persson discusses whether the average number of (different) countries listed on research reports is rising or declining. He seems to recognize that the ISI data published in our newsletter *Science Watch* (7 (1), 1; 1996) and picked up by your journal (*Nature* **379**, 287; 1996) deal with a different phenomenon: whether the percentage of internationally co-authored papers is markedly changing. (To clarify matters on this question, the *Science Watch* article found a "flattening out", elsewhere described as "a plateau". ISI found no evidence of a significant decline worldwide, or for any particular country, in percentage of international collaborations.) Thus, to characterize ISI's calculations as a miscount, when numbers collected and presented by Persson referred to a different question, is a mistake.

Furthermore, our analysis of Persson's question using a dataset larger than that used by Persson (22 OECD countries) — one incorporating all ISI-indexed journals in fact — does show the same flattening out in the average number of countries listed on research papers. The numbers ISI obtained, which are comparable to Persson's method of counting and are not limited to 22 OECD nations, are 1.124 for 1993 (938,482, the sum of countries on all papers, divided by 834,699, the number of papers) and 1.122 for 1994 (1,061,302 divided by 945,705). We would characterize this as a flattening out or a plateau.

David A. Pendlebury

Institute for Scientific Information,
Philadelphia,
Pennsylvania 19104, USA

1. *The Digest of Agricultural Census Statistics, United Kingdom, 1994*. (MAFF/HMSO, London, 1995).
2. Nix, J. *Farm Management Pocketbook, 26th edition* (1996). (Wye College Press, Wye, 1995).
3. Eurostat. *Agricultural Statistical Yearbook, 1995*. (Commission of the European Communities, Luxembourg, 1995).
4. Archer, J. *Crop Nutrition and Fertilizer Use*, 2nd Edn (Farming Press, Ipswich, 1988).
5. Howard, P. J. A. et al. *Soil Use Mgmt* **11**, 9–15 (1995).
6. Johnston, A. E. *Rothamsted Report for 1972, Part 2*, 131–159 (1973).
7. Johnston, A. E. In *Advances in Soil Organic Matter Research: The Impact on Agriculture and the Environment*. (ed. Wilson, W. S.) 299–314 (Royal Society of Chemistry, Cambridge, 1991).
8. Marland G., Andres, R. J. & Boden T. A. In *Trends '93: A Compendium of Data on Global Change*. (eds Boden, T. A. et al.) 505–584 (ORNL/CDIAC-65, Carbon Dioxide Information Analysis Center, Oak Ridge, 1994).
9. Whitmore, A. P., Bradbury, N. J. & Johnston, P. A. *Agric. Ecosys. Environ.* **39**, 221–233 (1991).

International collaboration

SIR — The Institute for Scientific Information (ISI) in Philadelphia recently reported in *Science Watch* (see *Nature* 379, 287; 1996) that “when one compares the percentage of internationally co-authored papers for 1992 against the figures for 1994, some of the countries have actually decreased their levels of international collaboration”. As well

Citation Index (SCI) data purchased on tape from ISI and thus excluding social science and arts and humanities publications. In addition, our analysis was restricted to articles, notes and review articles as these refereed publications most accurately reflect incremental contributions to the national and international innovation systems. Our results, therefore, are based on fewer papers; for example, in 1981 ISI counted 285,598 US papers while CHI Research, Inc. counted 134,769 US articles, notes and reviews in the natural and medical sciences (63% papers fewer than ISI).

The figure shows our findings. In 1981, of the 134,769 papers having a US author, 7.5% (10,116 papers) also had a foreign author. As the graph shows, this trend grew steadily to 16.1% in 1993. *Science Watch* reported a possible decline in US international collaborations. However, both the ISI and CHI data suggest we will have to wait to see if the decline is a trend or an artefact.

The evidence from SPRU research contradicts the UK trend reported in *Science Watch* and suggests that the growth

in international collaboration continues as reported earlier (see *Nature* 375, 99; 1995)². In 1981, of the 31,167 papers including a UK author, 13.7% (4,256 papers) also had a foreign author (4.6% with European Union (EU), 4.5% with US). *Science Watch* reported a plateau between 1992 and 1994 of about 22–23% of papers having an international collaboration. In contrast, we found that international collaboration with UK scientists grew steadily to 26.5% in 1994 (10.5% with EU, 8.3% with US).

These findings suggest that international collaboration in refereed contributions to the various innovation systems may still be increasing while international cooperation may be on the decline in non-refereed science, social science and arts and humanities publications. We wonder if the same applies to other G7 nations?

J. Sylvan Katz

Diana Hicks

*Science Policy Research Unit (SPRU),
University of Sussex,
Brighton, BN1 9RF, UK*

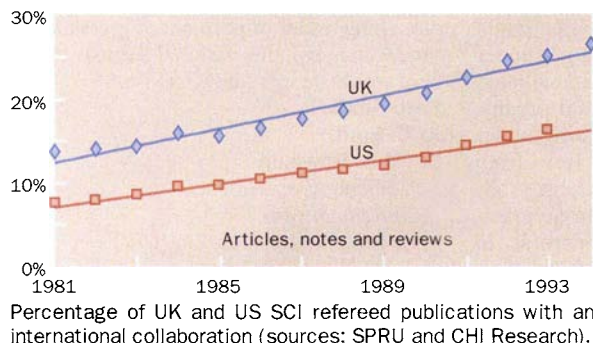
Francis Narin

Kimberly Hamilton

*CHI Research, Inc.,
10 White Horse Pike,
Haddon Heights,
New Jersey 08035, USA*

1. *Science Watch* 7 (1), 1996.

2. Katz J. S. et al. *Changing Shape of British Science* Appendix C, 1995 (SPRU, University of Sussex).



as scientific publications, *Science Watch* counted papers published in the social sciences and the arts and humanities indexed in their databases between 1981 and 1994. We wondered if declining trend in international collaboration would be found for more focused UK and US refereed papers in the natural and medical sciences.

We examined papers indexed in *Science*

The most direct
path to developing
purification
methods is via
AKTAexplorer

Access to data on humans

SIR — The announcement that *Nature* intends to impose public access to data on altered animals (*Nature* 379, 191; 1996) is welcome. But what about unaltered humans? Can we have your help in the acquisition and deposition of linkage data on which published papers are based? The data could be accessed by e-mail after publication. This would bring human data into line with the moral leadership long provided by the nematode project.

Although the human genome has vast resources devoted to its physical mapping, there are difficulties in acquiring good data on genetic mapping. Linkage studies are published without the requirement, usual in major journals for protein and DNA sequences, of easy access to raw data. The many successful studies in mendelian disorders have their data lost to posterity after a single pair of loci have defined a linkage, often after a trawl of hundreds of markers. Studies on the non-mendelian disorders usually have insufficient information to reach the arbitrary levels of likelihood demanded without the helping hand of chance, and their massive and reliable data are not available either for the general mapping on which related advances depend or as a contribution to specific mapping in the future. Few authors seem

aware of the necessary limitations of the present maps and the simple options available for their advance. As most linkage studies are analysed by a limited range of programs, compatibility is not a problem as the formats are well specified and simple.

The requirement of submission of data with manuscripts would have four advantages.

(1) It would allow data on genetic mapping to accumulate and be available for both global mapping and for scrutiny by those with special expertise in small segments of the genome.

(2) It would provide information on allelic frequencies and associations in different communities: this would give information on very close linkages as well as on the origins of populations and the mutation rates of the more mutable loci.

(3) It would allow the weak information on non-mendelian disorders, unbiased by selection for high lod scores at more than a few neighbouring loci, to be assembled together and analysed as a whole by the same programs, and also provide an incentive for program development. The vast diversion of blood and gold to linkage analysis in schizophrenia has so far led to only one hopeful segment with an allele sufficiently strong and frequent to be

noticed, although in no single study is the evidence convincing.

(4) It would allow reviewers the opportunity to examine the data, and their positive recommendations would be exposed to the quality control of independent review.

In practice, the development of DNA typing in humans has been associated with a degree of secrecy new to such studies and the ethical implications have been of little interest to those preoccupied with the expanding and invasive subject of medical ethics. The sheer size of the present genetic map imposes increasing restraints on the reliability of global ordering algorithms, especially where recombinant events cannot be counted, or are not counted, while the difficulty of the subject makes it impossible for anyone to be expert on more than a few segments of the order of a thousandth of our genome.

To exploit the genome project efficiently, we need to compare the structural or physical maps, progressing well in public documentation and methods of ordering and display, with the functional or genetic maps, on which raw data are limited and methods of ordering and display less advanced.

J. H. Edwards

*Department of Genetics,
University of Oxford,
Department of Genetics,
Oxford OX1 3QU, UK*

If you've ever developed a purification strategy by yourself, you know there is a lot to consider. Creating media screening schemes; designing buffer preparation routines; selecting which columns to use—even listing the tasks takes careful consideration. But now there's a better way of working.

Introducing ÄKTAexplorer

Turn ÄKTA™explorer on, choose a protocol, check running parameters and press start. ÄKTAexplorer does the rest.

Designed for all chromatographic techniques, this new purification system offers you pre-set protocols for every major purification task—including method scouting and media

screening. You'll save time as the system automatically recommends the best columns for your runs. You get fast pH screening as it automatically prepares your buffers from stock solutions. The moment you turn ÄKTAexplorer on, you're presented with a direct path to full-scale purification.

That path is UNICORN®—ÄKTAexplorer's control software. Your scale-up is simplified as the system's software can transfer and implement your methods on purification systems at all scales.

Of course, these are just a few of ÄKTAexplorer's features. So call us at 1 (800) 526 3593 from North America, +81 (0)3 3492 6949 from Japan or +46 (0)18 16 50 11 from the rest of the world (or meet us on the Internet at <http://www.biotech.pharmacia.se/akta.htm>).



**Pharmacia
Biotech**

Uppsala, Sweden. (And the rest of the world)

William Woodville and vaccination

SIR — Two hundred years ago, on 14 May 1796, Edward Jenner (1749–1823) inoculated a boy called James Phipps with a 'lymph' derived from a cowpox (vaccinia) vesicle affecting a dairymaid, Sarah Nelmes¹; more importantly, on 1 July, he took the adventurous (experimental) step of injecting him with smallpox 'matter', and Phipps did not develop variola (smallpox).

It is important however, to give due credit for widespread introduction of vaccination to William Woodville (1752–1805), Physician and Superintendent of the St Pancras Smallpox Hospital² in London, who had over many years taken a keen interest in all matters relating to smallpox prevention and treatment. (He had produced an impressive overview³ in 1796.)

Jenner was better known (and had been elected a Fellow of the Royal Society) for his seminal paper "Observations on the Natural History of the Cuckoo"⁴. His subsequent cowpox manuscript had been rejected for publication by Sir Joseph Banks, the President of the Royal Society, and Sir Everard Home; Jenner published it in expanded form as a pamphlet⁵.

Woodville did not immediately attempt confirmation of Jenner's claim — initially because of adverse reaction from the medical profession, and later because Jenner's supply of cowpox 'lymph' had been lost^{1,6}. Fortunately, he became aware of an outbreak of cowpox at Thomas Tanner's farm in Gray's Inn Lane in January 1799 (ref. 2); by comparison with coloured plates in Jenner's *Inquiry*⁵, Woodville (and several distinguished colleagues, including Banks) were able to satisfy themselves of the correct diagnosis. Woodville was thus able to vaccinate a large number of local residents, some of whom (almost certainly as a result of contamination of 'lymph' with variolous material) developed a generalized (vaccinal) rash. What was in fact the first large-scale clinical trial of vaccination confirmed Jenner's observation(s).

Woodville (at heart a scientist with a major interest in botany) was loth to admit contamination of his cowpox 'lymph'².

Nevertheless, although, as a result, relations between him and Jenner became cool⁶, the two physicians continued to communicate, remained on reasonably good terms, and subsequently published as a joint pamphlet *A Comparative Statement of Facts and Observations Relative to the Cow-Pox*⁷. In a dedication to Jenner in a subsequent pamphlet, *Observations on the Cow-pox*, Woodville wrote⁸: "That the vaccine matter, with which the inoculations have been carried on in the Hospital, was contaminated with that of the Variolous... is a charge which I know to be unfounded..." He continued: "The performance of this task [that is, refutation] has, however, been very painful to me... which attaches to a man, for whom I have long entertained a friendly regard..." Woodville subsequently played a major role in the introduction of vaccination in France^{1,6}; the technique was to be rapidly acclaimed internationally, although global eradication did not occur until October 1977.

Retrospectively, the fact that Woodville failed to receive due recognition for his role in the smallpox saga seems to have been largely because of his close association with George Pearson (1751–1828), a distinguished scientist who among other discoveries had recognized uric acid⁶. Pearson was exceedingly ambitious and ruthless; not only did he attempt to take personal credit for the demonstration of vaccination, but he did everything in his power to destroy Jenner's claim(s) for priority in its discovery. In fact, he also spoke (unsuccessfully) against the petition(s) to reward Jenner with parliamentary funds to the tune of £30,000 — a fortune indeed in 1800 (refs 1,6).

Woodville should therefore be accorded far greater credit for establishing vaccination on a large scale as a safe and effective protective measure against variola². Had it not been for his initiative and industry, Jenner's *Inquiry* might well have been neglected and his two subsequent pamphlets (of 1799 and 1800) remained unwritten. Jenner's 'thorn in the flesh' was certainly not the academic figure Woodville, but Pearson.

G. C. Cook

University College London Hospitals,
Hospital for Tropical Diseases,
4 St Pancras Way,
London NW1 0PE, UK

■ See also News and Views, page 26. □

Overdue changes

SIR — As a former Soros Visiting Scholar, I have first-hand experience of the importance of Mr George Soros's support for Russian science (see *Nature* 378, 432; 1995 & 380, 18; 1996). The scholarship gave me timely encouragement, and eventually

enabled me to raise funding for a doctorate at Oxford. On my return to Russia, however, I found that my research institute was not prepared to recognize my new degree. (Ironically, the director of the institute was himself recently awarded a Soros Professorship Award.)

My experience underlines the extent to which, despite the positive effects, the Soros scholarships scheme has failed to bring about the necessary structural changes in Russian science. Instead, it perpetuates the Stalin-designed system which isolates the research institutions of the Russian (formerly Soviet) Academy of Sciences from the universities that do the teaching. The academy carried out research either to meet the needs of the government, or to demonstrate the generosity of the Communist Party. In contrast, the universities have been expected to meet the manpower needs of industry, agriculture, education — and, in some cases, the academy itself.

Recent political developments have caused the government to lose interest in research, primarily because investment in basic research does not produce short-term cash returns, while the results of such research end up in the public domain. Universities — whose students, as the main consumers of basic research, could benefit from receiving this knowledge directly from research scientists — have not been able to link up with the academy's research institutions because of the restrictions imposed by the Ministry of Higher Education. At the same time, the universities lack the means to carry out research effectively.

Some universities have admittedly hired leading research scientists. But this is because of a shortage of teachers rather than a deliberate policy. The general difficulty faced by universities in attracting leading researchers reduces their attraction to students, and thus their government subsidies and fees. Members of the Academy of Sciences appear to have convinced Soros that, by helping them as individuals, he is rescuing Russian science. But they have not been interested in the structural changes that are sorely needed.

Such changes are now long overdue. No country can afford to support as many research institutions as Russia without linking research and higher education, and attracting funds from both government and external sources. Russia needs research universities — not just isolated and xenophobic research institutions.

Eugene Potapov

Institute of Biological
Problems of the North,
Russian Academy of Sciences,
Magadan 685000,
Russia

Present address: Department of Zoology, Uppsala University, Villavägen, S-752 36 Uppsala, Sweden.

1. Dixon, C. W. *Smallpox* 512 (J. & A. Churchill, London, 1962).

2. Cook, G. C. *J. med. Biogr.* (in the press).

3. Woodville, W. *The History of the Inoculation of the Smallpox in Great Britain; comprehending a review of all the publications on the subject: with an experimental inquiry into the relative advantages of every measure which has been deemed necessary in the process of inoculation* (James Phillips, London, 1796).

4. Jenner, E. *Phil. Trans. R. Soc.* 78, 219–237 (1788).

5. Jenner, E. *An Inquiry into the Causes and Effects of the Variolae Vaccinae, a disease discovered in some of the Western counties of England, especially Gloucestershire, and known by the name of cow pox*, 75 (Sampson Low, London, 1798).

6. Abrahams, J. J. *Lettsom: His life, times, friends and descendants*, 498 (Heinemann, London, 1933).

7. Jenner, E. & Woodville, W. *A Comparative Statement of Facts and Observations relative to the Cow-Pox*, 43 (Sampson Low, London, 1800).

8. Woodville, W. *Observations on the Cow-pox*, 43 (William Phillips, London, 1800).

Revealing the patterns of macroevolution

Robert L. Carroll

The evolution of amphibians from fish shows that major evolutionary transitions can occur through a step-by-step addition of novel characters over several million years.

ONE of the most critical problems in evolutionary theory has been to explain how totally new structures and ways of life have evolved. Many people — even many biologists — find it difficult to appreciate how small genetic differences between individual organisms, and their differential spread in populations due to natural selection, can explain the origin of characters such as wings in birds and flippers in whales, and the achievement of the distinct way of life that these structures imply.

When Simpson^{1,2} wrote his classic books on macroevolution (the large-scale patterns of evolutionary change), no major evolutionary transitions were adequately documented from the fossil record. He coined the term 'quantum evolution' to describe periods of apparently very rapid change for which there were very few intermediate fossils. He attributed the rapidity of change to the difficulty of adapting to conditions intermediate between major habitats or ways of life. Factors other than natural selection at the population level have also been proposed by Wright³, Gould⁴ and Stanley⁵, to explain the apparent rapidity of such transitions.

The transition between fish and amphibians (tetrapods) during the Late Devonian, some 365 million years ago, was one of the major events in the history of vertebrates. Changes in structure, physiology and behaviour were required for animals previously dependent on the water for support, locomotion, respiration and feeding to adapt to life on land. Here too it is difficult to explain how intermediate structures could have undergone gradual progressive adaptation to a succession of intermediate habitats; and, until recently, no fossils were known of intermediate forms.

On page 61 of this issue⁶, Ahlberg, Clack and Lukševičs provide evidence that helps to document the pattern and timescale of change from obligatorily aquatic fish to land vertebrates. Although only a few intermediates are known, individual genera show a mosaic pattern of primitive (old) and derived (newly evolved) characters that indicate different rates of evolution. Their paper deals specifically with differences in the degree of achievement of tetrapod characteristics in the skull roof, braincase and fins in *Pan-*

derichthys, a lobe-finned fish close to the ancestry of amphibians. This genus, from the late Middle or early Upper Devonian, has a larger number of characters in common with amphibians than any other lobe-finned fish (of which the coelacanth and lungfish are the sole survivors today).

Panderichthys has lost both dorsal and anal fins and evolved a slender tail. In contrast, the paired fins show no characteristics of land vertebrates. The skull roof

begun before the appearance of *Panderichthys*, more than 377 million years ago, and continued through early tetrapods such as *Acanthostega* (Figs 1, 2) at the end of the Devonian, some 14 million years later. In contrast, changes in the structure of the limbs and the braincase were not initiated in *Panderichthys*, but must have occurred later at a more rapid rate. Very primitive limb bones were described in *Elginerpeton* by Ahlberg⁷, from

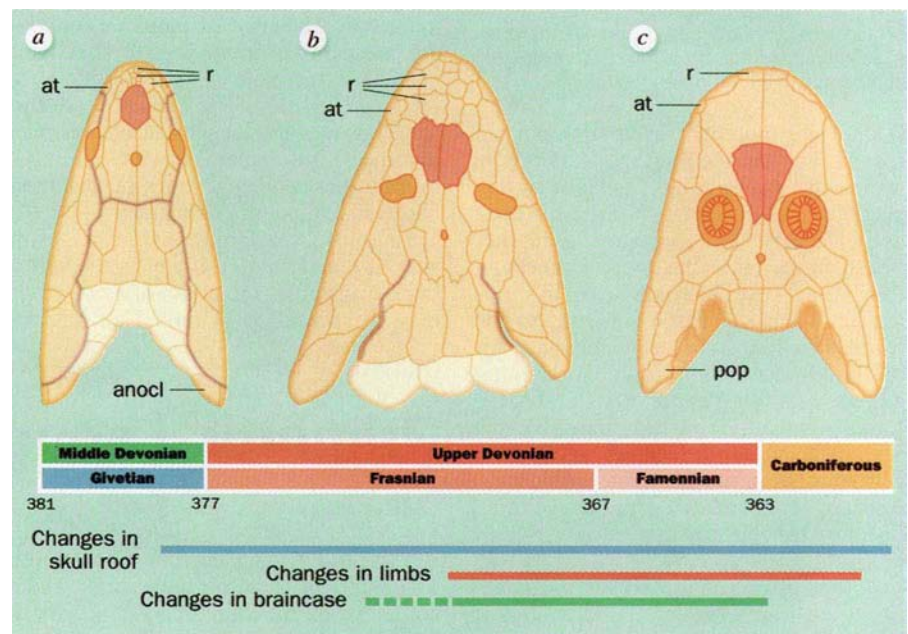


FIG. 1 Changes in the structure of the skull roof seen in the transition from fish to amphibians. *a*, *Eusthenopteron*, a typical lobe-finned fish from the early Upper Devonian. *b*, *Panderichthys*, an 'intermediate' form from the late Middle Devonian or early Upper Devonian. *c*, *Acanthostega*, an amphibian from the latest Devonian. Dark brown lines indicate mobility between areas of the skull roof that became progressively consolidated from the Middle to the uppermost Devonian. Paler bones are those lost in this transition. The bone in pink, the frontal, became paired and enlarged. Bones identified as at, r, anocl and pop are retained in Devonian amphibians, but lost in those of the Early Carboniferous. The mosaic pattern of changes in the skull roof, limbs and braincase are charted against the geological time scale (millions of years ago).

is better consolidated than that of more primitive genera in forming interlocking sutures between the bones of the snout and at the back of the skull table, and resembles that of early amphibians in the large paired frontal bones and in the position of the eyes, which have moved upwards and backwards (Fig. 1). In contrast, the braincase shows none of the derived characters of Upper Devonian tetrapods, but remains nearly identical with that of the typical lobe-finned fish, *Eusthenopteron*.

Changes in the skull roof must have

368-million-year-old deposits. Much more advanced limbs and/or girdles are evident in *Hynierpeton*, at approximately 366 million years, and in the Upper Devonian amphibians⁸⁻¹⁰. The major transformation of the limbs thus occurred rapidly over a period of about nine million years.

Ahlberg *et al.* argue that changes in the braincase occurred over approximately the same time span. The braincase in *Panderichthys* was composed of two elements that articulated loosely with one another and with the outer bones of the



FIG. 2 Articulated remains of the skeleton of the Upper Devonian amphibian *Acanthostega*, which provides the most complete evidence of the structure of early land vertebrates. (Geological Museum, Univ. Copenhagen; photograph by M. I. Coates.)

skull. Their mobility played an important role in aquatic feeding and respiration. Thomson¹¹ showed that elongation of that part of the skull lying in front of the eyes, as in *Panderichthys* and early tetrapods, would have made this system mechanically inoperative. In amphibians, in which the head was not firmly attached to the trunk, it would have been necessary for the bones of the skull to be closely integrated with one another to resist the forces of gravity and movements of the skull associated with feeding. So it is clear why changes in the braincase would have been closely correlated with the evolution of limbs capable of supporting the body on land.

Although only a few fossils document the transition between strictly aquatic lobe-finned fish and four-legged amphibians, they are enough to show that changes in the limbs and skull roof took place over nine to 14 million years, which is comparable with the time span of evolution in the size of the brain and the

pattern of locomotion that distinguish humans from apes. The transition between fish and amphibians is typically considered as the period during which paired fins evolved into limbs capable of terrestrial locomotion. But this is only part of a much longer sequence of changes in limb structure that began in the Late Silurian, about 405 million years ago, with the divergence of lobe-finned from more primitive ray-finned fish, and continued in the Lower Devonian with the initiation of the dichotomous structure of the lower limb bones common to *Eusthenopteron* and land vertebrates. Although a limb structure capable of terrestrial locomotion was achieved by the Late Devonian, this pattern has been continuously modified up to the present day.

Similarly, changes in the structure of the skull that were emphasized as part of the fish-amphibian transition actually began far earlier, and are still going on. Neither patterns nor rates of changes that are documented by Ahlberg *et al.* require evolutionary mechanisms beyond those that have been used to explain evolution within well-established groups. The rates are faster in some lineages and for some structures than in others, but remain several orders of magnitude slower than rates of evolution observed in modern populations¹².

As one of the largest-scale changes in structure and habitat of any of the vertebrate transitions, the origin of amphibians may be taken as a model of macro-

evolution. It is closely paralleled by the transition between terrestrial dinosaurs, through *Archaeopteryx*, to the much more advanced birds of the Lower Cretaceous (148–125 million years ago) recently outlined by Chiappe¹³. The structure and function of the individual flight feathers and their distribution on the wing had already achieved a fully modern appearance in *Archaeopteryx*, but the skeleton shows very few avian characters. Most of the modern features of the shoulder girdle and vertebral column were achieved within 10 to 15 million years after *Archaeopteryx*, but in a cumulative fashion. Changes in the wrist and hand, however, proceeded more slowly, and were not completed until several million years later. The loss of teeth apparently did not occur in the lineage leading to modern birds until near the end of the Cretaceous.

As transitions of roughly equivalent magnitude, the rates and patterns of changes are very similar to those observed in the origin of amphibians. In both cases, the fossil record demonstrates that major evolutionary transitions are accomplished through a succession of small changes, comparable to the pattern of evolution hypothesized by Darwin, although at a somewhat faster rate. □

Robert L. Carroll is at the Redpath Museum, McGill University, 859 Sherbrooke Street West, Montréal, Québec, Canada H3A 2K6.

ORIGINS OF LIFE

Primordial soup or crêpes?

Günter von Kiedrowski

ACCORDING to the current picture, life originated in a primordial ocean which contained all the ingredients necessary to form long, information-carrying polymers able to self-replicate, mutate and evolve. The recipe for a life-giving prebiotic broth was conceived by Darwin, refined by Oparin and Haldane, and finally elaborated by a whole community of prebiotic chemists starting with Miller in 1953. Although the taste of the soup seems quite agreeable now — at least with respect to its broad experimental base — there are critics who go for other courses as far as the origin of life is concerned. One of the oldest, yet still applicable, arguments against the soup is based on the thermodynamics and kinetics of polycondensation in aqueous solution. Hydrolysis will always limit chain length and prevent the formation of longer polymers that are necessary to set up a genetic system.

On page 59 of this issue¹, Ferris *et al.* provide evidence that longer oligonucleotides and peptides can be obtained if the polycondensation takes place on a

mineral surface instead of in free solution. For example, the polycondensation of ImpA, an activated nucleotide monomer, is rather inefficient in solution, barely yielding products longer than pentamers even in the presence of polyuridylic acid acting as a template^{2,3}. But when decaadenylic acid adsorbed on a montmorillonite clay is used as a primer, and the elongation reaction is repeated several times, oligomers up to 50-mers accumulate on the surface. This result is remarkable because it indicates that the clay surface may promote the formation of oligonucleotides up to the length of small ribozymes. The latter are being discussed in the context of the 'RNA world', a model of early life in which RNA molecules act both as enzymes and substrates for their own replication.

Peptide formation also benefits from surface chemistry. The authors show that polymerization of glutamic acid on illite (in the presence of a condensing agent) gives up to 55-mer-long products after 50 feedings with fresh reagents. Moreover,

1. Simpson, G. G. *Tempo and Mode in Evolution* (Columbia Univ. Press, New York, 1944).
2. Simpson, G. G. *The Major Features of Evolution* (Columbia Univ. Press, New York, 1953).
3. Wright, S. *Genetics* **16**, 97–159 (1931).
4. Gould, S. J. in *Tempo and Mode in Evolution* (eds Fitch, W. M. & Ayala, F. J.) 125–144 (National Academy of Sciences, Washington DC, 1995).
5. Stanley, S. M. *Proc. natn. Acad. Sci. U.S.A.* **72**, 646–650 (1975).
6. Ahlberg, P. E., Clack, J. A. & Lukševičs, E. *Nature* **381**, 61–64 (1996).
7. Ahlberg, P. E. *Nature* **373**, 420–425 (1995).
8. Coates, M. I. & Clack, J. A. *Bull. Mus. natl Hist. nat. Paris* **17**, 373–388 (1995).
9. Daeschler, E. B., Shubin, N. H., Thomson, K. S. & Amaral, W. W. *Science* **265**, 639–642 (1994).
10. Lebedev, O. A. & Coates, M. I. *Zool. J. Linn. Soc.* **114**, 307–348 (1995).
11. Thomson, K. S. *Zool. J. Linn. Soc.* **46**, 223–253 (1967).
12. Gingerich, P. D. *Science* **222**, 159–161 (1983).
13. Chiappe, L. *Nature* **378**, 349–355 (1995).

aspartic acid reacts on the surface of hydroxyapatite and in the presence of a water-soluble carbodiimide to give a material of such high molecular weight that the products cannot be separated chromatographically.

Taken together, these examples allow the fairly general conclusion that the polymers of life were more likely to have been baked like prebiotic crêpes than cooked in a prebiotic soup. French crêpes are prepared by pouring liquid dough over a hot stone plate, causing the dough to dehydrate and solidify. In the chemical analogy, dehydration corresponds to polycondensation, regardless of the required temperature. The physical base for the new recipe is similar to the one conceived by Bernal⁵, in which minerals absorb and concentrate the organic building blocks of life. In the picture put forward by Ferris and colleagues, minerals also bind the reaction products, negatively charged polyanions, mainly by electrostatic interactions. Binding increases with the length of the polymer. Above a critical chain length, polymer desorption is virtually impossible and any chain-elongating condensation step is forced to occur at the surface. This is the prebiotic equivalent of the solid-phase synthetic methods used to make nucleic-acid oligomers in biotechnology.

So prebiotic chemistry may now enjoy all the advantages that the use of a solid support gives to the organic synthesis of biopolymers. Binding and supporting oligomers, however, may not be the only benefit of a surface. At least in the case of clays, surface catalysis must be considered as the other essential. Based on detailed kinetic studies, Kawamura and Ferris have already discussed a model of clay catalysis in the oligomerization of activated nucleotides⁶. Moreover, the analogy with the conventional solid-phase synthesis should not be taken too strictly, as water must be avoided in these reactions. The average chain length is approximately given by the square-root of the ratio between the timescales of hydrolysis and of elongation. To obtain a chain length of 50, for example, elongation must occur 2,500 times faster than hydrolysis. This restricts the solid-phase scenario to relatively stable classes of polymers. Although reasonable figures on prebiotic polycondensation rates are hard

to estimate, peptides and nucleic acids are to be included owing to their slow hydrolysis.

In any case, the survival of a polymer may be easier when it is bound to a surface. A number of years ago, Günter Wächtershäuser introduced his 'chemo-autotrophic' theory of surface metabolism, on the basis of some general considerations of the thermodynamics and kinetics of bond breaking and making at surfaces and in solution⁷. Owing to their reduced translational and rotational freedom, surface-bound molecules are entropically more favoured to condense than those in solution. The restriction of movement makes condensation at surfaces more selective and perhaps somewhat slower than condensation in solution.

These arguments are clearly aimed at

burying the soup theory, but their significance to the above findings cannot be neglected. Does the present work point to the end of the soup era? Probably not. In Ferris and colleagues' model the soup is still required to provide the building material. It is thus compatible with a surface-mediated origin of life. Regardless of the particular model, the common message is that the earliest forms of life may have proliferated by spreading on surfaces. As far as information-carrying polymers are concerned, the mechanisms for template replication at surfaces clearly need to be elaborated. □

Günter von Kiedrowski is in the Lehrstuhl für Organische Chemie I, Bioorganische Chemie, Ruhr-Universität, 44780 Bochum, Germany.

IMMUNOLOGY

The two faces of the mast cell

Stephen J. Galli and Barry K. Wershil

No one who has seen (or survived) an anaphylactic reaction is likely ever to forget it^{1,2}. The cost of this exquisitely specific but catastrophic immune response, which can be triggered in sensitized individuals by such otherwise minor provocations as a bee sting or dose of penicillin, so grotesquely outweighs any apparent benefit that immunologists have long sought counterbalancing adaptive functions for one of the main culprits, the mast cell. On pages 75 and 77 of this issue, Echtenacher *et al.*³ and Malaviya *et al.*⁴ now report that the cell which is potentially deadly in anaphylaxis can also orchestrate life-saving host responses to bacterial infection.

This partial rehabilitation of the mast cell's reputation has been rather slow in coming. Over a century ago, Paul Ehrlich proposed that these cells, which he termed 'Mastzellen' (well-fed cells) because their cytoplasm is stuffed with prominent granules, may help to maintain the nutrition of connective tissues⁵, while Elie Metchnikoff was probably the first to suggest that mast cells had a phagocytic function, and might thereby contribute to host defence⁶. Although the involvement of mast cells in phagocytosis may be minor, at least in comparison to that of macrophages and neutrophils, there is no

question that mast cells are exceedingly good at initiating inflammation. They can bind immunoglobulin E (IgE) antibodies to receptors on their surface: aggregation of the receptors by contact of cell-bound IgE with multivalent antigen then leads to the rapid extracellular release of histamine, heparin, proteases and other mediators that are stored in the cells' cytoplasmic granules, and the synthesis and secretion of leukotrienes and prostaglandins. These products result in bronchoconstriction, changes in blood-vessel tone, increased vascular permeabil-

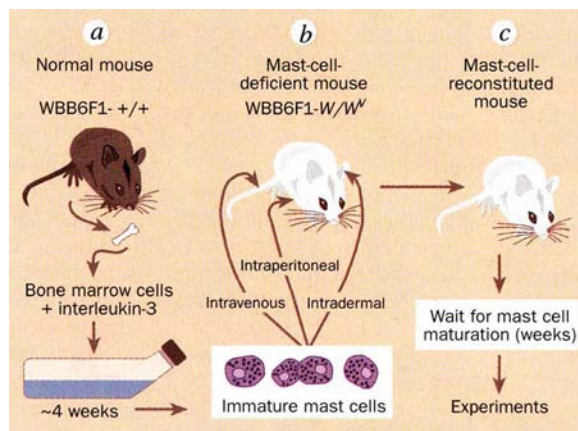


FIG. 1 Reconstituting mast-cell activity in mice. *a*, Generate lineage-committed, immature mast cells *in vitro*, from bone-marrow cells derived from normal mice that are congenic to WBB6F1-W/W^v mast-cell-deficient mice (or, for certain mechanistic studies, from mice with defined mutations that affect mast-cell development, mediator content or function). *b*, Transfer immature mast cells intravenously, intraperitoneally or into specific tissues (for instance the skin or gastrointestinal tract) of genetically mast-cell-deficient W/W^v mice. *c*, Wait several weeks to allow mast cells to mature and acquire phenotypic characteristics that are appropriate for the sites of transfer.

1. Ferris, J. P., Hill, A. R. Jr, Liu, R. & Orgel, L. E. *Nature* **381**, 59–61 (1996).
2. Sawai, H. & Orgel, L. E. *J. Am. chem. Soc.* **97**, 3532–3533 (1975).
3. Sleeper, H. L., Lohrmann, R. & Orgel, L. E. *J. molec. Evol.* **13**, 203–214 (1979).
4. Szostak, J. W. & Ellington, A. D. in *The RNA World* (eds Gesteland, R. F. & Atkins, J. F.) 511–533 (Cold Spring Harbor Laboratory Press, NY, 1993).
5. Bernal, J. D. *The Physical Base of Life* (Routledge & Kegan Paul, London, 1951).
6. Kawamura, K. & Ferris, J. P. *J. Am. chem. Soc.* **116**, 7564–7572 (1994).
7. Wächtershäuser, G. *Microbiol. Rev.* **52**, 452–484 (1988).

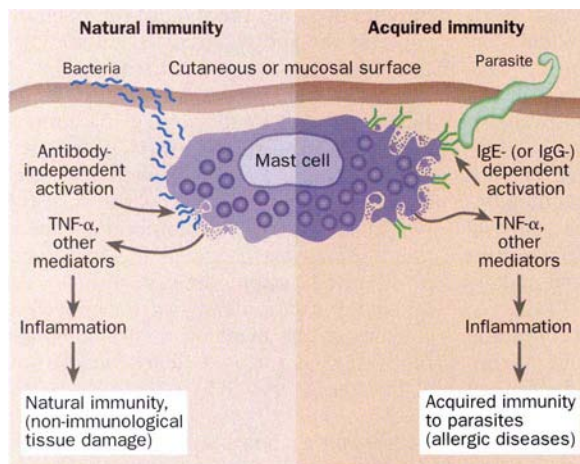


FIG. 2 Mast cells can function in host defence (or disease) by releasing $\text{TNF-}\alpha$ and other mediators that orchestrate local inflammatory responses. Mast-cell mediators released by immunologically nonspecific mechanisms can participate in natural immunity (against bacteria, for example), but may also induce non-immunological tissue damage; depending on the specificity of their cell-bound antibodies, mast cells can also take part in acquired immune responses (against parasites, for example) or contribute to the pathogenesis of allergic diseases.

ity, and myriad other pro-inflammatory effects^{1,7,8}.

Yet the question remained: although mast cells clearly have the means and the opportunity to initiate IgE (or IgG) antibody-dependent anaphylactic reactions, as well as to contribute to the pathophysiology of asthma and other allergic disorders, what is the motive? Why would nature design an element of acquired immunity to be so potent, and so widely distributed in vascularized tissues, that its activation in response to intrinsically innocuous substances can result in the death of the host? One possibility is that the mast cell has adaptive functions when, in localized microenvironments or sub-maximal amounts, it releases many of the same mediators that are produced widely, rapidly and in large quantities during anaphylaxis.

But immune responses may involve the coordinated and potentially redundant activities of several cell types — so how can one characterize the specific contributions of a single player? In the case of the mast cell, this problem can be addressed by using mice in which mast-cell activity has been selectively reconstituted (ref. 9; Fig. 1) to compare the expression of biological responses in tissues that differ solely in containing, or virtually lacking, mast cells. In a separate development, mast cells were revealed to be potential sources of many cytokines, at least one of which, tumour necrosis factor- α ($\text{TNF-}\alpha$), can be released rapidly from preformed stores and, with more prolonged kinetics, from transcripts produced in response to appropriate activation of the cell^{10,11}. Mice with reconstituted mast cells were then used to show that $\text{TNF-}\alpha$ is an important

mediator of mast-cell-dependent leukocyte recruitment, both in response to IgE and antigen¹² or immune complexes¹³, and that mast cells are required for the expression of IgE-dependent host defence against at least one species of ectoparasite¹⁴.

Should one then believe that the mast cell has become (at least in the developed world) the 'appendix' of the human immune system — that is, a cell which had an essential function in acquired immunity in our more parasite-infested past but which is now mainly an effector cell in allergic diseases? Probably not. Acute inflammation has long been regarded as a critical element of intrinsic or 'natural' immunity, and, under some circumstances,

bacterial products (such as lipopolysaccharide) and other non-immunological stimuli can promote mediator release from mast cells^{1,8,10,15}. Moreover, mast cells can markedly amplify many features of immunologically non-specific acute inflammatory responses, including neutrophil recruitment¹⁶.

The two new papers demonstrate that mast cells represent a central component of host defence against bacterial infection, and indicate that mast-cell- and $\text{TNF-}\alpha$ -dependent recruitment of circulating leukocytes with bactericidal properties is crucial in this response. Malaviya *et al.*⁴ show that mast cells can promote the clearance of virulent strains of *Klebsiella pneumoniae* from the lungs or the peritoneal cavity of mice, as well as reduce the mortality associated with experimental intraperitoneal infection; and they provide evidence that $\text{TNF-}\alpha$ released by mast cells, apparently in response to contact with a component of the bacteria (type I fimbriae), is important in the generation of the associated influx of neutrophils.

Echtenacher *et al.*³ report that reconstitution of mast-cell-deficient W/W^v mice with mast cells also prevents death from surgically induced bacterial peritonitis, and that this effect can be abolished by the administration of antibodies to $\text{TNF-}\alpha$. The injection of W/W^v mice with $\text{TNF-}\alpha$ (instead of mast cells) is also protective, but only when the cytokine is provided in the right amount. Again, one apparently can have either too little or too much of a good thing!

In all, these various lines of research suggest that mast cells lead a double life in host defence: they are important com-

ponents of natural immunity as well as critical effectors in acquired immunity (Fig. 2). Both Metchnikoff and Ehrlich would be pleased.

Yet even if one is interested only in the involvement of the mast cell in host defence (which is much too narrow a concept of its potential functions), much work remains to be done. How broad, and how important, is the mast cell's role in natural or acquired responses to different types of pathogens; what signals regulate mast-cell mediator release in these settings; are there quantitative or qualitative differences in the mast-cell mediators released during natural or acquired immune responses; do mast cells have other important functions in these responses (such as regulating the epithelial barrier, or contributing to antigen presentation or immunoregulation), as well as initiating inflammation; to what extent do the findings in mice apply to humans; and can the function of mast cells in host defence be manipulated for therapeutic ends?

Despite these uncertainties, it seems clear that, depending on the circumstances, the release of $\text{TNF-}\alpha$ and other mediators by mast cells can indeed confer either great benefit or harm. It appears that nature has distributed large numbers of mast cells near surfaces exposed to the external environment, and designed them to release their mediators rapidly — not to permit mast cells to precipitate anaphylaxis or the inflammation associated with allergic diseases, but so that they can help to orchestrate protective host inflammatory responses against pathogens. □

Stephen J. Galli and Barry K. Wershil are at Beth Israel Hospital, 330 Brookline Avenue, Research North, Boston, Massachusetts 02215, USA.

- Bochner, B. S. & Lichtenstein, L. M. *New Engl. J. Med.* **324**, 1785–1790 (1991).
- Colten, H. R. *Nature* **371**, 474–475 (1994).
- Echtenacher, B., Männel, D. N. & Hülthner, L. *Nature* **381**, 75–77 (1996).
- Malaviya, R., Ikeda, T., Ross, E. & Abraham, S. N. *Nature* **381**, 77–80 (1996).
- Ehrlich, P. *Beiträge zur Theorie und Praxis der histologischen Färbung* (thesis, Univ. Leipzig, 1878).
- Metchnikoff, E. *Leçons sur la Pathologie Comparée de l'Inflammation* (Masson, Paris, 1892).
- Kinet, J.-P. & Metzger, H. in *Fc Receptors and the Actions of Antibodies* (ed. Metzger, H.) 239–259 (Am. Soc. Microbiol., Washington DC, 1990).
- Schwartz, L. & Austen, K. F. *Prog. Allergy* **34**, 271–321 (1984).
- Nakano, T. *et al. J. exp. Med.* **162**, 1025–1043 (1985).
- Young, J. D.-E., Liu, C.-C., Butler, G., Cohn, Z. A. & Galli, S. J. *Proc. natn. Acad. Sci. U.S.A.* **84**, 9175–9179 (1987).
- Gordon, J. R. & Galli, S. J. *Nature* **346**, 274–276 (1990).
- Wershil, B. K., Wang, Z.-S., Gordon, J. R. & Galli, S. J. *J. clin. Invest.* **87**, 446–453 (1991).
- Zhang, Y., Ramos, B. F. & Jakschik, B. A. *Science* **258**, 1957–1959 (1992).
- Matsuda, H. *et al. J. Immun.* **144**, 259–262 (1990).
- Leal-Berumen, I., Conlon, P. & Marshall, J. S. *J. Immun.* **152**, 5468–5476 (1994).
- Wershil, B. K., Murakami, T. & Galli, S. J. *J. Immun.* **140**, 2356–2360 (1988).

The mountains will flow

Philip England

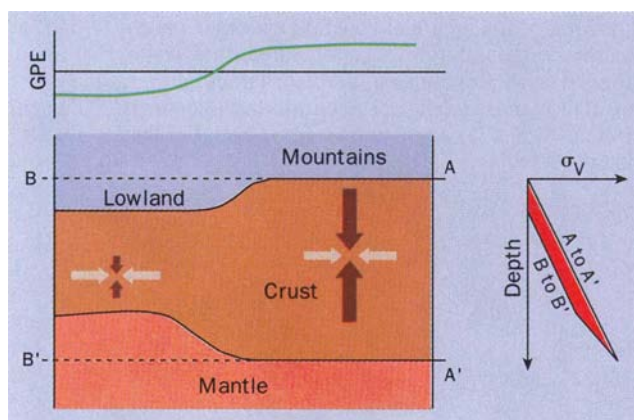
THE key to understanding many geological processes lies in knowing which parts of the solid Earth act as fluids over geological time spans (10^3 to 10^9 years). The fluidity of a system is expressed by its Deborah number, which is the ratio of the length of time needed for flow to occur to the timescale over which a force is applied¹. The eponymous Deborah prophesied that the mountains would flow before the Lord², but contemporary geological opinion is divided on whether mountains flow on a secular basis. On page 37 of this issue³, Jones *et al.* present evidence that the mountains of western North America, at least, are behaving as a fluid.

Following the success of plate tectonics in describing the oceanic portions of the Earth's surface as a set of rigid shells, the theory was applied to the continents with a flair and enthusiasm that, for a while, disguised its severe shortcomings as a description of continental deformation. By the late 1970s, however, it had become clear that the continents behave in a fundamentally different way from the oceans. Unlike the narrow boundaries a few tens of kilometres wide between oceanic plates, zones of deformation in the continents are hundreds to thousands of kilometres wide — larger than some of the plates^{4,5}. Therefore, the simplicity of plate tectonics, which derives from the rigidity of the plates, is absent from the continents. There is little agreement, however, as how best to describe this more complex deformation.

Some view mountain building as a process of viscous deformation, and others see it as the relative motion of many rigid micro-plates⁶. One observation offers the possibility of distinguishing between these views: the distribution of density inside the continents appears to influence their deformation. The continental crust floats on the denser mantle as icebergs float on water, with high mountains generally being underlain by thicker crust (see figure). This state of affairs is referred to by geologists as 'isostatic balance'.

Isostasy is a balance of vertical forces, but two isostatically balanced columns may be in different states of stress because of differences between their density structures or, equivalently, between their gravitational potential energy (GPE).

Broadly speaking, the average pressure is higher beneath mountains than beneath lowlands. We may expect that high mountains should extend under their own weight⁷ and, more generally, that movements within continental regions should be from high ground towards lowland⁸. This suggestion has qualitative support from the distribution of deformation in the eastern Mediterranean⁸, Asia⁹ and the Andes¹⁰, but a quantitative analysis



In isostatically balanced columns of crust, the vertical force σ_v on a horizontal plane is equal to the weight of overlying rock. This force is greater beneath mountains than beneath lower ground. The profiles of σ_v beneath a mountain range with high surface and thick crust (AA') and beneath thinner crust (BB') are shown schematically at right. The mountain may be stretching horizontally, while the lowland, with the same horizontal stress (white arrows), is shortening¹⁰. The difference in gravitational potential energy (GPE) between two regions is a measure of the difference in σ_v (red area at right).

offers the possibility of measuring the strength of the continents.

Such an analysis is provided by Jones and colleagues³, who exploit the dense distribution of geophysical data across western North America to investigate the relationship between the state of stress, implied by differences in GPE, and the rate of strain (deformation). They estimate GPE in two complementary ways. The first uses seismic wave speeds to estimate the subsurface density distribution. The second estimate comes from observations of the geoid, the equipotential surface of gravity that best fits mean sea level. Undulations of this surface carry

information about the subsurface density distribution. Each method requires assumptions, either to convert seismic wave speed into density or to remove from the geoid the influence of density distributions deep in the mantle. Either method alone might be suspect, but the concordance between the two suggests that they are reliable.

The observed strain rates in western North America show that the regions with higher GPE are extending, whereas those below a certain energy level are in compression. The story is not entirely simple, however. Some of the regions with the highest potential energy are not extending as rapidly as regions with lower GPE. Jones and co-workers are able to show that these observations are explained by differences in temperature: regions through which the surface heat flux is lower are deforming less rapidly than neighbouring, hotter, regions. This result is consistent with the general observation that rocks become weaker as their temperature is increased.

It seems that the distribution of present-day deformation within western North America is governed by forces that result from the density distribution inside the continent. This result is important for two reasons. First, the history of the deformation of North America is most commonly interpreted as the result of changes in the forces applied to its western edge due to changes in plate motion.

Although the plate boundary forces must have been important in building up the mountains of North America, Jones *et al.* show that the potential energy stored there is still dominating deformation 50 million years after the mountains were built. If interior forces are important today, then they were probably more important in the past, when the mountains were higher^{11,12}, so tectonic reconstructions based purely on relative plate motions, and ignoring interior forces, are surely flawed.

Second, the GPE differences within North America are similar in magnitude to those associated with density contrast

1. Reiner, M. *Deformation, Strain and Flow* (H. K. Lewis, London, 1969).

2. Judges 5: 5.

3. Jones, C. H., Unruh, J. R. & Sonder, L. J. *Nature* **381**, 37–41 (1996).

4. Molnar, P. & Tapponnier, P. *Science* **199**, 419–426 (1975).

5. McKenzie, D. *Geophys. J. R. astr. Soc.* **55**, 217–254 (1978).

6. Molnar, P. *Nature* **335**, 131–137 (1988).

7. Molnar, P. & Lyon-Caen, H. in *Processes in Continental and Lithospheric Deformation* (ed. Clark, S. P. Jr) (Geol. Soc. Am. spec. Pap. 218, 1988).

8. McKenzie, D. P. *Geophys. J. R. astr. Soc.* **30**, 109–185 (1972).

9. Molnar, P., England, P. & Martinod, J. *Rev. Geophys.* **31**, 357–396 (1993).

10. Dalmayrac, B. & Molnar, P. *Earth planet. Sci. Lett.* **55**, 473–481 (1981).

11. Molnar, P. & Chen, W.-P. *J. geophys. Res.* **88**, 1180–1196 (1983).

12. Sonder, L. J., England, P. C., Wernicke, B. P. & Christiansen, R. L. in *Continental Extensional Tectonics* (eds Coward, M. P., Dewey, J. F. & Hancock, P. L.) 187–201 (Geol. Soc. spec. Publ. No. 28, 1987).

within the oceanic plates. Yet the plates remain essentially undeformed over life-spans of 100 to 200 million years, whereas North America is straining at a rate of 10^{-15} to 10^{-16} s^{-1} in response to forces of the same magnitude. These rates of strain may seem slow, but they are brisk in geological terms — they

would produce 100 per cent strain in 20 to 200 million years.

The continent of North America must be much weaker than the stiff oceanic plates, so weak in fact that it makes no sense to treat it as a plate at all. Mountains do indeed flow, though at a rate that is slow on a human timescale. Deborah, I

presume, intended her prophecy as a comment on the Lord's omnipotence; it could equally be interpreted as a comment on His longevity. □

Philip England is in the Department of Earth Sciences, University of Oxford, Parks Road, Oxford OX1 3PR, UK.

OBITUARY

Colin S. Pittendrigh (1918–96)

COLIN S. Pittendrigh, founder of the modern field of circadian biology and for many years its most articulate spokesman, died in Bozeman, Montana, on 18 March.

Pitt, as he was universally known, took a first degree in botany at the University of Durham, and in the Second World War worked on malaria control in Trinidad, studying the anopheline mosquitoes that breed in bromeliad ponds. During that period he made substantial contributions to understanding of bromeliad evolution and his interest in rhythmicity was kindled by observing the biting rhythms of malaria-carrying mosquitoes. After the war he completed a PhD under the supervision of Theodosius Dobzhansky at Columbia University, moving to Princeton in 1947. There he taught a memorable course in introductory biology, and notably, as dean of the graduate school, was instrumental in making admission to Princeton open to women. In 1969 he moved to Stanford, where he remained until his retirement in 1984.

His early and seminal insight was that the many rhythms expressed by living things in the real world are the consequence of endogenous, self-sustained biological oscillations that are entrained (synchronized) by environmental cycles. Although this view seems obvious now, it was accepted only slowly because it seemed unnecessarily complex. Why should natural selection favour an internal oscillator when the environment is replete with temporal cues? One answer was contained in the earlier research of G. Kramer and K. von Frisch on Sun-compass orientation in birds and bees. To use the Sun as a compass animals need the equivalent of a clock, and it was this work that led Pitt to the inspired guess that this same clock times many biological processes and is indeed a general property of cells and organisms.

Many clear predictions emerged from this new view of rhythmicity. In the mid-1950s, Pitt, along with his graduate students and close associate Victor Bruce, began testing them. Pitt himself concentrated on the beautifully precise

eclosion (pupal emergence) rhythm of *Drosophila pseudoobscura* and quickly verified three central predictions: the eclosion rhythm persists in the absence of temporal cues from the environment (that is, under constant conditions in the laboratory); its period varies only slightly with changes in ambient temperature (a prerequisite for accurate time-keeping); and it can be entrained by cycles of light



or temperature within a narrow range around its natural period.

Together with his long-time friend and scientific colleague, Jürgen Aschoff, Pitt led the work of the next decade which established that biological clocks with nearly identical formal properties occur in virtually all eukaryotic organisms. They provide temporal frameworks for processes as different as photosynthesis in green plants, spore formation in fungi, cell division in the ears of mice, and human cognitive performance. The recent demonstration of a clock in a cyanobacterium extends the generalization to the prokaryotes.

Although Pitt participated in one of the earliest identifications of a circadian pacemaker in a multicellular organism (in the optic lobe of the cockroach), physiology was not his *métier*. While

many in the field moved in that direction, he began to study photoperiodic time measurement, an interest that occupied the rest of his career. Building on a prescient insight of Erwin Bünning and a beautiful experiment by K. K. Nanda and K. C. Hamner, he developed a set of logically tight arguments and designed a series of elegant experiments that convincingly demonstrated that the circadian clock measures daylength and thus times many of the adaptive responses of organisms to seasonal change.

Pitt was always interested in the power of language to direct thought, and over the past few years he became convinced that the term 'biological clock' had outlived its usefulness. He perceived that the central function of the circadian system is to provide a temporal framework within the organism or cell: to emphasize this view, he preferred the term 'temporal programme' to 'clock'.

Although his experimental work was almost entirely directed at understanding circadian rhythmicity, his thinking was dominated by evolutionary considerations. In one of his papers — which is still well worth reading — he originated the concept of 'teleonomy' as a necessary and respectable alternative to 'teleology' with its unacceptable baggage (Chapter 8 in *Behavior and Evolution*, Yale University Press, 1958).

By force of intellect and personality, Colin Pittendrigh exerted a profound influence on the development of a field which, very early on, he realized is central to our understanding of biological organization. It has taken many years for biologists generally to catch up with his vision, but it is now clear that he was right and the hunt is on for underlying circadian mechanisms at the cellular and molecular levels. Those of us — and there are many — who were privileged to know and work with Pitt will never forget him as a scientist or as a man.

Michael Menaker

Michael Menaker is in the Department of Biology, University of Virginia, Charlottesville, Virginia 22903-2477, USA.

Proton decay at the drip-line

Philip Woods

WHAT determines whether a nuclear species exists? For nuclear scientists the answer to this poser is that the species should live long enough to be identified and its properties studied. This still begs the fundamental scientific question of what the ultimate boundaries to nuclear existence are. Work by Davids *et al.*¹ at Argonne National Laboratory in the United States has pinpointed the remotest border post to date, with the discovery of proton decay from ¹⁸⁵Bi, which has 24 neutrons fewer than the only stable isotope of bismuth, ²⁰⁹Bi.

Nuclear proton decay occurs when there is a severe deficit of neutrons compared with stable nuclear species. Neutrons provide the glue that prevents the nucleus blowing itself apart under the mutual electrostatic repulsion of the protons. If too many neutrons are removed from a nucleus it eventually becomes energetically unstable to the spontaneous emission of a proton, as nature protests against the extremes to which it is being pushed. This ultimate limit to nuclear stability is known as the proton drip-line.

The time taken for a proton to cross the nucleus is $\sim 10^{-22}$ s, and state-of-the-art techniques for detecting nuclear proton decay are limited to lifetimes in excess of 10^{-7} s, so one might ask how it is possible to study this phenomenon. In the nuclear surface region, the combination of the attractive strong nuclear force and the repulsive electrostatic force gives rise to a potential energy barrier that the proton must penetrate for the decay to occur. If a nucleus has a large number of protons (the atomic number, Z) then this barrier can become high enough for the proton to be restrained inside it long enough for the isotope to be observed.

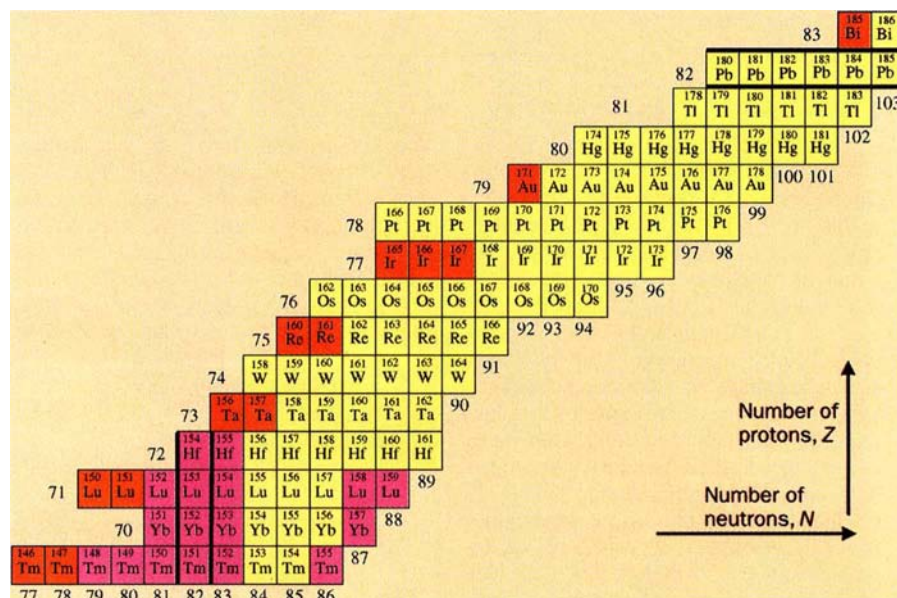
Proton decay therefore represents a beautifully simple example of a quantum tunnelling process in which, before emerging, the proton must traverse ~ 100 femtometres ($\sim 10^{-13}$ m) of the zone classically forbidden because its energy is too low. Ironically, alpha-decay — a more familiar and common nuclear decay mechanism — is inherently a much more complex process, involving the formation of ⁴He clusters in the nuclear surface region. The simplicity of the proton decay mechanism, requiring only the emission of a single constituent nucleon, can be exploited to provide a unique window on the properties of nuclei at the extreme edge of existence. The potential energy barrier is critically influenced by the centrifugal potential experienced by the proton, which makes proton decay rates extremely sensitive to the orbital angular momentum of the proton in the host nucleus. Hence pro-

ton decay half-life measurements can be used to explore nuclear shell structures in this twilight zone of nuclear existence.

The proton drip-line sounds an interesting place to visit. So how do we get there? The answer is an old one: fusion. In this case, the fusion of heavy nuclei produces highly neutron-deficient compound nuclei, which rapidly de-excite by boiling off particles and gamma-rays, leaving behind a plethora of highly unstable nuclear species. Only a small fraction of these, about 10^{-5} to 10^{-7} , may be the

case since it has one proton more than the magic number $Z=82$, corresponding to the closure of a major nuclear shell. In fact the observed proton decay comes from an excited intruder state configuration formed by the promotion of a proton from below the $Z=82$ shell closure. This decay transition was used to provide unique information on the quantum mixing between normal and intruder state configurations in the daughter nucleus.

Following the serendipitous discovery of nuclear proton decay⁴ from an excited state in 1970, and the first example of ground-state proton decay⁵ in 1981, there followed relative lulls in activity. Now we are entering a renaissance of proton drip-line research, in which we can aim to fully



The proton drip-line, as currently established, between thulium and bismuth. The red isotopes decay by proton emission, yellow by α and β only, and pink by β only. All the proton emitters have odd Z , corresponding to an unpaired, energetically unbound proton. Proton emission from even- Z isotopes has yet to be identified, because attractive pairing interactions push the drip-line beyond current isotope production limits. In future it is hoped to observe two-proton emission from even- Z nuclei, in cases where single-proton emission is energetically forbidden.

ton-unstable isotope of interest. Ultra-sensitive techniques, involving in-flight electromagnetic separation devices and sophisticated detection systems in which radioactive ions are implanted before decaying, have been developed to filter out and identify these extremely rare isotopes. Consequently, we can now explore extended regions of the proton drip-line in detail for the first time.

Particularly dramatic progress has been made in the region of odd- Z elements between thulium ($Z=69$) and gold ($Z=79$) for which a continuous series of proton-decaying isotopes has been discovered^{2,3}. Most recently, work by Davids *et al.* has established the current extreme outpost for such studies with the identification of proton emission from the isotope ¹⁸⁵Bi, which represents by far the heaviest point at which the drip-line has been crossed. This isotope is a particularly interesting

map one of the fundamental limits to nuclear stability. Furthermore, new techniques have been developed to study in detail gamma-ray de-excitation cascades from excited proton-radioactive nuclei⁶. With the advent of radioactive beam facilities allied to these techniques, the prospects for nuclear drip-line research look highly promising. □

Philip Woods is in the Department of Physics and Astronomy, University of Edinburgh, Mayfield Road, Edinburgh EH9 3JZ, UK.

1. Davids, C. N. *et al.* *Phys. Rev. Lett.* **76**, 592–595 (1996).
2. Woods, P. J. *et al.* *Nucl. Phys. A* **553**, 485–488 (1993).
3. Davids, C. N. *et al.* *Proc. Exotic Nuclei and Atomic Masses Conf.*, Arles, 1995 (in the press).
4. Jackson, K. P. *et al.* *Phys. Lett.* **B33**, 281–284 (1970).
5. Hofmann, S. *et al.* *Proc. 4th International Conf. Nuclei far from Stability* CERN 81-09, 190–201 (1981).
6. Paul, E. S. *et al.* *Phys. Rev.* **C51**, 78–87 (1995).

Protease inhibitor implicated

James O. McNamara and Ram S. Puranam

THE culprit responsible for a rare form of inherited epilepsy, known as progressive myoclonic epilepsy of the Unverricht-Lundborg type (EPM1), has been colared. Writing in *Science*¹, Pennacchio *et al.* describe how they have exposed the identity of the mutant gene that is carried unbeknown by both parents and can manifest as EPM1 epilepsy in their offspring. The culprit, surprisingly, had not even been a suspect — of all the aberrant proteins that might have been expected to be encoded by this gene, which include ones affecting neuronal excitability or synaptic

transmission, this turns out to be a cysteine protease inhibitor, whose normal function is to prevent the destruction of key proteins.

The epilepsies comprise a common collection of disorders that affect at least one per cent of the population worldwide and which consist of more than 40 distinct types, all sharing the common feature of an enduring increase of neuronal excitability that causes a periodic and unpredictable occurrence of seizures. EPM1 is an exceedingly rare form. Children afflicted with it start to suffer seizures

between the ages of 8 and 13 (ref. 2). There follows a gradual decline in neurological function and the intellect deteriorates slowly, with a loss of roughly ten IQ points per decade. Survival into adulthood is common and some individuals reach the sixth decade. Autopsy discloses widespread neuronal degeneration.

The past decade has witnessed enormous progress in our understanding of the mechanisms that underlie a number of these disorders. But the restraint imposed by the daunting complexity of the mammalian nervous system makes the power of genetics especially appealing. Many of the epilepsies recognized to have genetic determinants (about one-third of cases) appear to be complex genetic disorders where several genes are probably at fault,

MEDICAL HISTORY

Country doctor and speckled monster

As the world knows, Edward Jenner (1749–1823) was the man who, through his observation of the immunity of milkmaids in his native Gloucestershire to the 'speckled monster', as smallpox was then known, established the means in 1796 by which the disease was eventually to be eradicated. A statue to commemorate him has long stood in Kensington Gardens, London; but it had become forgotten, and last week a ceremony was held there as a feature of the bicentennial celebrations of the advent of vaccination.

Jenner has sat there reflectively, chin in hand, for nearly 150 years. The sculptor, William Calder Marshall, gave a clue by inscribing a head of a cow on the chair on which his model sat. But only the name — Jenner — was inscribed on the pedestal and so, over time, the memorial became virtually forgotten. The Friends of Hyde Park and Kensington Gardens 'rediscovered' the memorial in their effort to foster recognition of the unique Victorian statues in Kensington Gardens. They were then joined by the Jenner Educational Trust and St George's Hospital Medical School, and together they have now placed a commemorative plaque at the memorial — at last giving full recognition to Jenner.

It is difficult to imagine the horror that smallpox brought in its train. No one was safe. In Jenner's time, it caused ten per cent of the total deaths and a third of those of children in London. It was the cruellest of the emissaries of the Grim Reaper.

On 14 May 1796, lymph from a pustule of dairymaid Sarah Nelmes was vaccinated by Jenner into James Phipps. Later, when the boy was inoculated with smallpox, the feared symptoms — high fever, followed by spots and, often, death — failed to appear.

Vaccination was shown to give protection. It then took nearly 200 years before the World Health Organization, on completing an international eradication programme, declared in 1980 that the world was rid of the disease.

A number of statues of Jenner have been erected. The first was placed in

that vaccination gave her family.

But Jenner's statue in Kensington Gardens is his main memorial. Calder Marshall was a competent artist of the Victorian era and it was he, in fact, who first conceived the idea of the memorial. The necessary funds were slow in coming until an appeal was launched internationally. The United States headed the donations; Russia, despite the intervention of the Crimean War, came second; and Britain only third.

With the permission of Queen Victoria, a site in Trafalgar Square was secured and Prince Albert, the Prince Consort, presided over an inaugural occasion in 1858. Everybody who was anybody attended. But, soon afterwards, Jenner was banished. A non-military character was thought inappropriate in an area devoted to British success at arms. *The Times* spoke up for his removal, and it was demanded in Parliament. *Punch* took an ironic view:

*England's ingratitude still blots
The escutcheon of the brave and free;
I saved you many million spots,
And now you grudge one spot for me.*

And so, in 1862, Jenner was moved to become the first statue in Kensington Gardens. St George's Hospital, originally at Hyde Park Corner, put in a bid for the statue on the centennial anniversary of his discoveries. It had good grounds: the illustrious John Hunter, a surgeon at the hospital, was Jenner's mentor. But the attempt failed.

And so, today, his memorial still stands in Kensington Gardens. But now with acknowledgement of his place in medical history.

John Empson

John Empson is at 26 Pembroke Place, Kensington, London W8 6EU, UK.

■ See also page 18.

Dan Atkin

Jenner reflects in Kensington Gardens.

1825 in Gloucester Cathedral. Another, erected in 1865, stands in Bologne-sur-Mer, where the first vaccinations in France were carried out. Later, a memorial was placed in 1904 in Tokyo in the gardens of the National Museum. Perhaps the best known, with a cast in the Wellcome Building in London, is in the Museum of Fine Art in Genoa. It was by a leading Italian sculptor, Giulio Monteverde, and was commissioned in 1873 by Maria Brignole Sale, Duchess of Galliera, in recognition of the protection

although a tiny fraction (less than one per cent of cases) do seem to be accounted for by a single mutant gene.

Single mutant genes have been identified with three of the monogenic human epilepsies. A point mutation (caused by a base substitution in the DNA) in a gene encoding the $\alpha 4$ subunit of the neuronal nicotinic acetylcholine receptor underlies one form of frontal lobe epilepsy³. The other two monogenic epilepsies are like EPM1 in that they are myoclonic, with characteristic violent muscle contractions: a point mutation in a mitochondrial gene encoding a lysyl transfer RNA is responsible for myoclonic epilepsy and ragged red fibre disease⁴, and deletions and a point mutation that alters a splice site in a novel gene have been linked to Batten disease⁵.

Pennacchio *et al.*¹ also used a genetic approach to determine the nature of the defective protein in EPM1 epilepsy. Linkage analysis had already localized the gene underlying EPM1 to a region of 2 million base pairs on the long arm of chromosome 21, which was eventually narrowed down to 175 kilobases. Using direct complementary DNA selection, the authors looked for expressed DNA within this region of the genome: one cDNA was found to encode cystatin B, a protein inhibitor of cysteine proteases. This gene was expressed in all tissues examined.

Lymphoblastoid cell lines from affected and unaffected individuals were used to compare the expression of cystatin B messenger RNA. Affected individuals from four families had undetectable levels of cystatin B mRNA, unlike asymptomatic carriers and unaffected individuals. This suggested that a mutation was present in the cystatin B gene that resulted in decreased amounts of the mRNA.

Sequencing the gene from affected individuals disclosed different mutations in different families. A point mutation found in affected members of one family lay in a 3' splice site, a G-to-C transversion at the last nucleotide of intron I, a position highly conserved in all introns. A second mutation, at amino-acid position 68, that generated a stop codon, was found in two additional families. Oddly, no mutation could be identified in the fourth family, despite sequencing the entire coding region and large tracts of non-coding regions. The authors conclude that there is a strong association between the symptoms suffered by patients with EPM1 (the phenotype) and the abnormal reduction of cystatin B mRNA, although a causal link still needs to be proven.

An odd feature peculiar to EPM1 that may provide a clue to understanding some of the phenotypic consequences of this mutation is the deleterious effect of the antiseizure drug phenytoin on these patients⁶. Phenytoin is widely used and is safe and effective in many patients with epilepsy, but not for those with EPM1, in

whom it speeds neurological deterioration and may even shorten lifespan.

How might a mutation leading to reduced cystatin B mRNA (and presumably cystatin B protein) produce an epileptic phenotype? Cysteine proteases are small proteins that catalyse the destruction of diverse polypeptide substrates and are most active under mild reducing or acidic conditions, properties that render them well adapted to the environment of lysosomes⁷, which are organelles inside the cell where unwanted macromolecules are broken down. These enzymes can work on a broad range of substrates and are important in regulating intracellular protein turnover.

The cystatins are naturally occurring protein inhibitors of cysteine proteases and include more than 20 related members of a superfamily; cystatin B is itself a member of a subgroup of small intracellular proteins. It is widely distributed in cells and tissues, including the brain. This ubiquitous distribution, together with its apparent locale in the cytosol, has led to the suggestion⁸ that cystatin B protects the cell from cysteine proteases leaking out of the lysosomes.

The discovery by Pennacchio and colleagues¹ of the association of EPM1 with inadequate amounts of cystatin B mRNA raises many questions. Assuming that the phenotype is due to defective cystatin B activity intrinsic to the brain, in which brain regions is cystatin B expressed? Is it expressed in neurons and/or glia? Is it restricted to the cytosolic compartment of

these cells? How is the expression of cystatin B and its target proteins regulated during development? What is the key enzyme(s) inhibited by cystatin B specifically in the brain? Alternatively, does cystatin B have functions apart from being a protease inhibitor?

The answers to these questions, together with the creation of mice carrying null mutations of cystatin B, should help us decipher how this mutation produces the EPM1 phenotype. Insight derived from the study of EPM1 may provide a clue to the cellular and molecular mechanisms underlying some of the more common forms of human epilepsy with similar features, and lead to more effective therapy. □

James O. McNamara is in the Departments of Medicine (Neurology), Neurobiology and Pharmacology, and Ram S. Puranam is in the Department of Medicine (Neurology), Duke University and Durham Veterans Affairs Medical Centers, Durham, North Carolina 27710, USA.

1. Pennacchio, L. A. *et al.* *Science* **271**, 1731-1734 (1996).
2. Berkovic, S. F., Andermann, F., Carpenter, S. & Wolfe, L. S. *New Engl. J. Med.* **315**, 296-305 (1986).
3. Steinlein, O. K. *et al.* *Nature Genet.* **11**, 201-203 (1995).
4. Shoffner, J. M. *et al.* *Cell* **61**, 931-937 (1990).
5. International Batten Disease Consortium *Cell* **82**, 949-957 (1995).
6. Eldridge, R., Stern, R., Iivanainen, M., Koerber, T. & Wilder, B. J. *Lancet* **II**, 838-842 (1983).
7. Henskens, Y. M. C., Veerman, E. C. I. & Amerongen, A. V. N. *Hoppe-Seyler Z. biol. Chem.* **377**, 71-86 (1996).
8. Barrett, A. J. *Ann. N. Y. Acad. Sci.* **674**, 1-15 (1992).

OPTICS

Light bent by magnets

G. W. 't Hooft and M. B. van der Mark

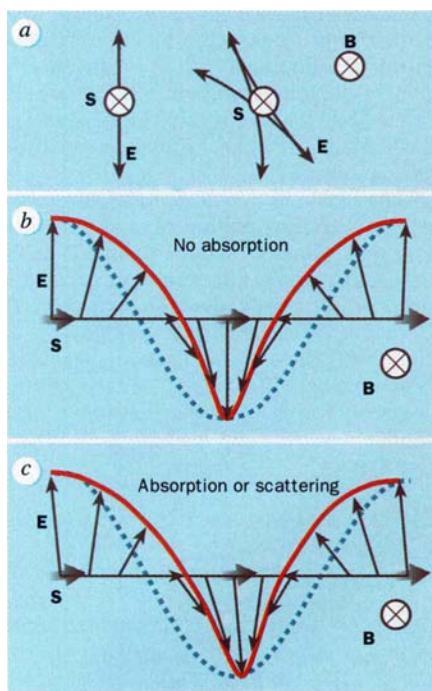
ON page 54 of this issue¹, Rikken and van Tiggelen report a new optical effect. In their experiment, light is injected into an isotropic scattering medium in a direction perpendicular to an applied magnetic field. They see a light flux induced in the third direction, perpendicular to the field and to the injected light, which is directly proportional to the magnetic field. Is this a trivial effect or is it surprisingly new?

An external magnetic field influences the refractive index of an isotropic dielectric medium. When the light travels parallel to the magnetic field, the refractive index for right circularly polarized light differs from the one for left circularly polarized light. This is known as the Faraday effect, and it is proportional to the magnetic field.

For linearly polarized light, the magnetic field causes the polarization plane to rotate in its passage through the medium. Part *a* of the figure shows this effect, where the electric field is repre-

sented by arrows and the wave vector and the external magnetic field are perpendicular to the plane of drawing. The charges in the dielectric are driven by the electric field of the light. Owing to the presence of an external magnetic field, the moving charges will experience a sideways Lorentz force. The net result is that the electric field vector is dragged along with the moving charges, thereby rotating. The Faraday effect cannot give rise to beam walk-off; in other words the Poynting vector, $\mathbf{S} = \mathbf{E} \times \mathbf{H}$, which describes the energy flow, will be parallel to the wave vector.

When we change the geometry so that the external magnetic field is perpendicular to the wave vector, the refractive index for light polarized parallel to the external field will differ from the one with perpendicular polarization. Only the latter case is of interest here and is shown in part *b* of the figure. By the reasoning above, the electric field vector will experience a wig-



The Faraday effect and beam walk-off. The thin arrows are the electric field vectors of light, $\mathbf{E}$, and the thick arrows show the Poynting vector $\mathbf{S}$ (denoting the direction of energy flow); an externally applied magnetic field $\mathbf{B}$ points into the plane of drawing in each case. *a*, Light moving parallel to $\mathbf{B}$ has its plane of polarization rotated. *b*, Moving perpendicular to $\mathbf{B}$, a longitudinal wobble is induced in the electric field. *c*, In an absorbing medium this produces a transverse flux of energy because the wobble and the main component of $\mathbf{E}$ are no longer perfectly out of phase.

gling in the plane of the drawing, perpendicular to the external magnetic field. So the Faraday effect induces an electric field in the direction of the wave, but one that is much smaller than the original perpendicular component.

The two components of the electric field are exactly out of phase with one another. The accompanying Poynting vector will also wiggle, but on average, over a full cycle, it will point in the same direction as the wave vector. In order for beam walk-off to occur, the medium must be absorptive, as in part *c*. Because of the presence of absorption the two components are no longer completely out of phase. At those positions where the amplitude of the electric field is largest (the positions which contribute the most to the average energy flow) the Poynting vector is not parallel to the wave vector. Now the external field points into the plane of drawing, the wave vector points to the right, and the Poynting vector has a deviation perpendicular to both.

The phenomenon observed by Rikken and van Tiggelen happens in an inhomogeneous, isotropic medium without absorption. In view of the above this seems quite surprising. The effect is a result of

the presence of scatterers². The similarity between absorption and scattering can be made plausible by the notion that both give rise to an exponentially decaying intensity for the directly transmitted part of the light.

For a high enough concentration of scatterers, the intensity distribution of the scattered light can be described by a diffusion equation. This is similar to the diffusion process of molecules in a gas, although in the case of photons the total number of particles is not constant. For a mixture of a polyatomic gas with a noble gas, transverse diffusion has been observed experimentally³. The explanation of this effect runs as follows: in a diffusing gas there is a particle flow in which non-spherical molecules line up to reduce their collision probability. This is analogous to the alignment of tree trunks thrown into a flowing river. The magnetic field then causes the magnetic moment of the molecules to rotate around the field. This precession will change the collision cross-section, and consequently alter the diffusive flow of the molecules. In the same way, stirring the river will decrease the throughput and will also make some of the trunks end up on the river banks.

In the case of the diffusion of light, an external magnetic field will change the refractive index of the scatterers and thereby alter the scattering process, which in turn will give rise to a different diffusion constant. Rotational symmetry around the external field axis and time-reversal symmetry arguments demand that changing the sign of the magnetic field reverses the transverse diffusion, as verified by Rikken and van Tiggelen¹.

The Faraday effect is large near an optical resonance and changes sign there, so two frequency components either side of a resonance should be partially separated. A further test would be to verify the change in the forward and backward light flux induced by a magnetic field.

It is too early to say whether there will be applications of the phenomena described above. One should bear in mind, however, that optical gain plays exactly the same symmetry-breaking role as absorption or scattering. Consequently, lasers in an external perpendicular magnetic field will also cause beam walk-off in the third direction, and here the effect may be a big one. □

G. W. 't Hooft and M. B. van der Mark are at the Philips Research Laboratories, Prof. Holstlaan 4, 5656 AA Eindhoven, The Netherlands.

1. Rikken, G. L. J. A. & van Tiggelen, B. A. *Nature* **381**, 54–55 (1996).
2. van Tiggelen, B. A. *Phys. Rev. Lett.* **75**, 422–424 (1995).
3. Mazur, E., Hijnen, H. J. M., Hermans, L. J. F. & Beenakker, J. J. M. *Physica* **123A**, 412–427 (1984).

Good neighbours

GROWING old is a melancholy fate. All the systems of the body start to 'go downhill'; joints degrade, arteries harden, tissue loses its elasticity, metabolism slows. But the body doesn't simply wear out like a bicycle or a car. It ages by the failure of its self-repair mechanism. Every now and again, some random error in the DNA of a cell evades this mechanism. Deprived of a gene, the cell becomes slightly less efficient. When it divides, its daughters inherit the inefficiency. If the error is in the DNA-repair machinery itself, that line of cells is likely to get very bad indeed. And less than 1 per cent of bad cells can be fatal. You will probably reach your deathbed with over 99 per cent of your cells in perfectly good shape.

Daedalus has a way out. His idea is for the many good cells to repair the few bad ones. Suppose, he says, that a defective cell could be fused to one of its neighbours. As in normal heredity, the defect will be recessive. The good cell will supply the gene missing from the bad one, and the fused product will function perfectly. Even if that gene is not missing, but present, damaged and actively mischievous, all is not lost. Let the bad cell be fused with two or three good neighbours, and they will outvote the aberration. The joint venture will still divide normally; it and its daughters will be good cells.

To work this trick, Daedalus is adapting his scheme of last week for fusing fat cells together in the body. This puts a high-frequency alternating field from a diathermy machine across the whole body, and concentrates its action in a small region by means of a phase-locked focused beam of ultrasound at the same frequency. This combination can fuse cells fairly efficiently. But how to target the few aberrant cells? Daedalus points out that they must differ slightly from their neighbours, in size, shape or conductivity. Electrical breakdown always concentrates at such anomalies and discontinuities. So with the right electric and ultrasonic intensities, Daedalus's gadget will concentrate its energy on the cellular aberrations. It will fuse them to their neighbours, while ignoring the mass of good cells. By sweeping the ultrasonic focus around the body, bad cells everywhere could be fused and reclaimed.

Thus the decrepitudes of age will be averted. A periodic 'sweep' of bad cells will keep you young, spry and active for ever — well, not quite ever. One day in late age, the cell-death programmed into your cells, at least the good ones, will begin to bite. On that day, only bad cells will survive. And there won't be enough of them.

David Jones

Interleukin-16 or not?

SIR — As part of an intensive search for natural suppressors of HIV replication produced by T lymphocytes, Baier *et al.*¹ reported a novel antiviral activity of interleukin-16 (IL-16), a CD8⁺ T-cell-derived cytokine. This molecule was first identified^{2,3} as a lymphocyte chemoattractant factor (LCF) because of its ability to promote the migration of CD4⁺ T cells. More recent studies⁴ indicate that this factor is released from CD8⁺ T cells after stimulation by antigens, histamine or mitogens; however, the deduced protein sequence of human LCF lacks a recognizable signal peptide to facilitate secretion⁵.

Puzzling observations^{1,5} about the true relative molecular mass (M_r) of the biologically active form(s) of LCF — the recombinant protein tends to aggregate^{1,5} and may have folding difficulties¹ — led us to examine the published LCF polypeptide chain sequence for structural information that might provide clues to this behaviour. Instead, we find peculiar, correlated anomalies in the nucleotide sequence and the derived protein sequence of LCF/IL-16 that are inconsistent with previous reports about the properties of this molecule.

We scanned the complementary DNA sequence of LCF and its predicted protein product⁵ against the nonredundant databases of the National Center for Biotechnology Information (NCBI), using the BLAST network server. The 130-amino-

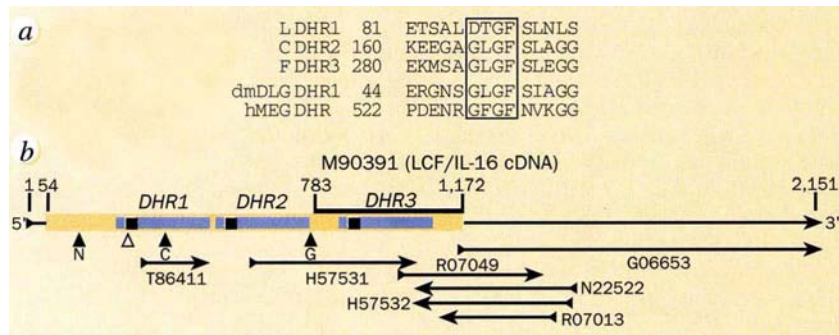
acid LCF protein chain retrieved a diverse group of intracellular proteins with a recurring module of about 80 residues called a 'discs-large homologous region' (DHR), named after repeats seen in the product of the *Drosophila dlg* tumour-suppressor gene^{6,7}; these DHR modules appear to direct specific protein-protein interactions inside the cell^{7,8}. Although the 14,000- M_r LCF protein has a single, embedded DHR motif⁷, we found two additional, deduced, DHR-like protein sequences from a region corresponding to the 5'-untranslated region of the gene preceding the LCF coding segment (see figure).

To explain this finding, we suggest that sequencing errors present in the original LCF cDNA may have produced frameshifts obscuring a longer protein-coding region. Support for this idea comes from a collection of expressed sequence tags (ESTs) from the NCBI database that coincidentally covers large tracts of the LCF cDNA and permits partial 'repair' of an extended coding region by direct alignment (see figure). We pinpointed potential frameshifts in the LCF cDNA by protein matches with DHR homologues that force jumps between translation reading frames⁹. In addition, we used a frameshift-error detection algorithm that tracks statistically significant, frame-dependent properties of coding regions¹⁰. Our results suggest that the restoration of

three principal missing bases (a C and a G after positions 400 and 780, respectively, and an unknown base after position 150 of the LCF cDNA) to the cDNA sequence results in a different LCF protein product of 374 amino acids ($M_r \sim 40,000$), containing three DHR repeats (see figure) and still no hydrophobic signal peptide. All other DHR-containing proteins are located inside the cell⁷.

This deduced LCF clone does not seem consistent with the original suggestion that a short fragment of its protein product functions as a classical, secreted cytokine^{2,3}. There are at least three possible explanations for this discrepancy. First, LCF activity as historically reported is due to a molecule not encoded by the cDNA in question. Indeed, the molecular identity of LCF may remain unclear: disparity between the cloned 14,000- M_r LCF protein⁵ and the initially characterized 50,000–60,000- M_r LCF activity^{2,3} has been attributed to a required autoaggregation of the smaller species^{1,5}. Second, LCF activity is perhaps embodied in the protein fragment encoded by the cDNA, but is of questionable physiological relevance. The defining biological measure of LCF is the ability to bind CD4 (as principally assessed by antibody-blocking studies^{5,11,12}); however, this observed affinity may not constitute a specific ligand-receptor interaction, and the observed movement of CD4⁺ cells *in vitro* does not necessarily translate to directed migration *in vivo*. For example, experiments using eosinophils show that diverse molecules that bind CD4 (such as OKT4 antibody and HIV envelope glycoprotein gp120) also induce cellular migration¹², but are effects that are probably not relevant *in vivo*. Finally, LCF activity may be authentic and perhaps physiologically relevant, but it is generated by processes atypical of most cytokines — for example, by proteolysis inside the cell and release by unknown mechanisms or on cell lysis.

In sum, the activities associated with LCF/IL-16 could result from the fortuitous association of the truncated



Reconstructing the IL-16/LCF coding region. **a**, Sequence alignment of the DHR-diagnostic Gly-Leu-Gly-Phe (GLGF) sequence patterns^{6,7} from the three deduced 40,000- M_r LCF DHR modules (numbered 1–3) with two typical DHR domains from *Drosophila* DlgA (dmDLG DHR1) and human MEG tyrosine phosphatase (hMEG DHR)⁷; the attached number locates the first amino acid of the sequence; the central GLGF motifs are boxed. The DHR3-equivalent sequence of LCF is noted in ref. 7; DHR modules have been redesignated as PDZ domains¹³. **b**, The complete, 2,151-bp cDNA isolated for human LCF/IL-16 (ref. 5) (Genbank accession number M90391) aligned to EST fragments T86411, H57531, R07049, N22522, H57532 and R07013, and the STS sequence G06653. Black triangles, potential frameshift locations (and deduced, missing bases) in the LCF cDNA indicated by the overlapping EST/STS sequences and the frameshift detection algorithm; open triangle, short region around positions 270–300 that may locally shift to a neighbouring frame. The resulting open reading frame (yellow box; bp 54–1,172; N-terminal amino acids Met-Pro-Ser-Gln...) encodes a 374-amino-acid protein composed primarily of three DHR modules^{6,7} (blue boxes) labelled as above; black squares, GLGF patterns; overhead bar, fragment composing the published LCF/IL-16 sequence⁵ (bp 783–1,172) containing DHR3.

1. Baier, M. *et al.* *Nature* **378**, 563 (1995).
2. Center, D. M. & Cruikshank, W. J. *Immun.* **128**, 2563–2568 (1982).
3. Cruikshank, W. & Center, D. M. *J. Immun.* **128**, 2569–2574 (1982).
4. Laberge, S. *et al.* *J. Immun.* **156**, 310–315 (1996).
5. Cruikshank, W. W. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **91**, 5109–5113 (1994).
6. Woods, D. F. & Bryant, P. J. *Mech. Dev.* **44**, 85–89 (1993).
7. Ponting, C. P. & Phillips, C. *Trends biochem. Sci.* **20**, 102–103 (1995).
8. Sato, T. *et al.* *Science* **268**, 411–415 (1995).
9. Claverie, J.-M. *J. molec. Biol.* **234**, 1140–1157 (1993).
10. Fichant, G. A. & Quentin, Y. *Nucleic Acids Res.* **23**, 2900–2908 (1995).
11. Cruikshank, W. *et al.* *J. Immun.* **138**, 3817–3823 (1987).
12. Rand, T. H. *et al.* *J. exp. Med.* **173**, 1521–1528 (1991).
13. Kennedy, M. B. *Trends biochem. Sci.* **20**, 350 (1995).
14. Kozak, M. *J. Cell Biol.* **108**, 229–241 (1989).
15. Cho, K.-O. *et al.* *Neuron* **9**, 929–942 (1992).
16. Laberge, S. *et al.* *J. Immun.* **155**, 2902–2910 (1995).

CCT	CGA	GAG	CTG	TCA	ACA	CAG	GCT	GAG	GAA	TCT	CAA	GGC	CCA	GTG	CTC	AAG	ATG	CCT	AGC	60
Pro	Arg	Glu	Leu	Ser	Thr	Gln	Ala	Glu	Glu	Ser	Gln	Gly	Pro	Val	Leu	Lys	Met	Pro	Ser	120
CGA	CGA	GAG	CTG	TCA	ACA	CAG	GCT	GAG	GAA	TCT	CAA	GGC	CCA	GTG	CTC	AAG	ATG	CCT	AGC	120
Gln	Arg	Ala	Arg	Ser	Phe	Pro	Leu	Thr	Arg	Ser	Gln	Ser	Gln	Ser	Gln	Ser	Gln	Ser	Gln	180
GAA	AAG	ACC	AGC	AAA	CTC	TAT	TCT	ATC	AGC	AGC	CAA	GTG	TCA	TGG	GCT	GTC	ATG	AAA	TCC	240
Glu	Lys	Thr	Ser	Lys	Leu	Tyr	Ser	Ile	Ser	Ser	Gln	Val	Ser	Ser	Ala	Val	Met	Lys	Ser	300
TTG	TGC	TCT	CTC	ATC	TCT	ATC	TCT	ATC	TCT	ATC	TCT	ATC	TCT	ATC	TCT	ATC	TCT	ATC	TCT	360
Leu	Leu	Cys	Leu	Pro	Ser	Ser	Ile	Ser	Ser	Cys	Ala	Gln	Thr	Pro	Cys	Ile	Pro	Lys	Ser	420
GCA	TCT	CCA	ACA	TCA	TCA	TCC	AAC	GAA	GAC	TCA	GCT	GCA	AAT	GCT	TCT	GCT	GAA	ACA	CTC	480
Ala	Ser	Pro	Thr	Ser	Ser	Ser	Asn	Glu	Asp	Ser	Ala	Ala	Asn	Gly	Ser	Ala	Glu	Thr	Ser	540
GCC	TTG	GAC	ACA	GGG	TTC	TGC	TGC	AAC	CTT	TCA	GAG	CTG	AGA	GAA	TAT	ACA	GAG	GCT	GTC	600
Ala	Leu	Asp	Thr	Gly	Phe	Ser	Leu	Asn	Leu	Ser	Glu	Leu	Arg	Glu	Tyr	Thr	Glu	Gly	Thr	660
ACG	GAA	GCC	AAG	GAA	GAC	GAT	GAT	GGG	GAC	CAC	AGT	TCC	CTT	CAG	TCT	GCT	CAC	TCC	GTT	720
Thr	Glu	Ala	Lys	Glu	Asp	Asp	Asp	Gly	Asp	His	Ser	Ser	Leu	Gln	Ser	Gly	Gln	Ser	Val	780
ATC	TCC	CTG	CTG	AGC	TCA	GAA	GAA	TTA	AAA	AAA	CTC	ATC	GAG	GAG	GTG	AGG	GTT	CTG	GAT	840
Ile	Ser	Leu	Leu	Ser	Ser	Glu	Glu	Leu	Lys	Lys	Leu	Ile	Ala	Glu	Val	Lys	Val	Leu	Asp	900
GAA	GCA	ACA	TTA	AAG	CAA	TTA	GAC	GGC	ATC	CAT	GTC	ATC	TTA	CAC	GAG	GAA	GAG	GAA	GAT	960
Glu	Ala	Thr	Leu	Lys	Gln	Leu	Asp	Gly	Ile	His	Val	Thr	Ile	Leu	His	Lys	Glu	Gly	Thr	1020
GCT	GAT	CTT	GGG	TTC	AGC	TTG	GCA	GGA	GCA	GAT	CTA	GAA	AAC	GAG	ATC	ATT	ACG	GTT	ACT	1080
Ala	Gly	Leu	Gly	Phe	Ser	Leu	Ala	Gly	Gly	Ala	Asp	Leu	Glu	Asn	Lys	Val	Ile	Thr	Val	1140
ACA	AGA	GTG	TTT	CCA	AAT	GGG	CTG	GGC	TCC	CAC	GAA	GGG	ACT	ATT	CAG	AAG	GGC	AAT	GAG	1200
His	Arg	Val	Phe	Pro	Asn	Gly	Leu	Ala	Ser	Gln	Glu	Gly	Thr	Ile	Gln	Lys	Gly	Asn	Glu	1260
GTT	CTT	TCC	ATC	AAC	GGC	AAG	TCT	CTC	AAG	GGG	ACC	ACG	CAC	CAT	GAT	GCC	TTG	GCA	ACT	1320
Val	Leu	Ser	Ile	Asn	Gly	Lys	Ser	Leu	Lys	Gly	Thr	His	His	Asp	Ala	Leu	Ala	Ile	Ala	1380
CTC	CGC	CAA	GCT	CGA	GCC	AGG	CAA	GCT	GTG	ATT	GTC	ACA	AGG	AAG	CTG	ATC	CCA	GAG	Lys	1440
Leu	Arg	Gln	Ala	Arg	Glu	Pro	Arg	Gln	Ala	Val	Ile	Val	Thr	Arg	Lys	Leu	Thr	Arg	Pro	1500
GCC	ATG	CCT	GAC	CTC	AAC	TCC	TCC	ACT	GAC	TCT	GCA	GCC	TCA	GCC	TCT	GCA	GCC	AGT	GAT	1560
Ala	Met	Pro	Asp	Leu	Asn	Ser	Ser	Thr	Asp	Ser	Ala	Ser	Ala	Ser	Ala	Ser	Ala	Ser	Ala	1620
GTT	TCT	GTG	GAA	TCT	ACA	GCA	GAG	GCC	ACA	GTG	ACG	GTG	ACA	CTG	GAG	AAG	ATG	TGC	CTC	1680
Val	Ser	Val	Glu	Ser	Thr	Ala	Glu	Ala	Thr	Val	Cys	Val	Thr	Leu	Glu	Lys	Met	Ser	Thr	1740
GCA	GGG	CTG	GGC	TTG	AGC	CTG	GAA	GGG	AGG	GCC	TCC	CTA	CAC	GAG	GAG	ACC	CTC	CTC	CTC	1800
Ala	Gly	Leu	Gly	Phe	Ser	Leu	Gly	Gly	Gly	Lys	Gly	Ser	Leu	His	Gly	Asp	Pro	Leu	Leu	1860
ACC	ATT	AAC	AGG	ATT	TTG	AAA	GGA	GCA	ACC	TCA	GAA	CAA	AGT	GAC	ACA	GCT	CGA	GCT	CTC	1920
Thr	Ile	Asn	Arg	Ile	Phe	Lys	Gly	Ala	Ala	Ser	Glu	Gln	Ser	Glu	Thr	Val	Gln	Pro	Gly	1980
GAT	GAA	ATC	TTG	CAG	CTG	GCT	GGT	CGC	ATC	GCC	ATG	CAG	GCC	CTC	ACA	CGG	TTT	GAA	GCC	2040
Asp	Glu	Ile	Leu	Gln	Leu	Gly	Gly	Thr	Ala	Met	Gln	Gly	Leu	Thr	Arg	Phe	Glu	Ala	TGG	2100
AAC	ATC	ATC	AAG	GCA	CTG	CCT	GAT	GGA	CCT	GTC	ACG	ATT	GTG	ATC	AGG	AGA	AAA	AGC	CTC	2160
Asn	Ile	Ile	Lys	Ala	Leu	Pro	Asp	Gly	Pro	Val	Thr	Ile	Val	Ile	Arg	Arg	Lys	Ser	Leu	2220
CAG	TCC	AAG	GAA	ACC	ACA	GCT	GCT	GGA	GAC	TCC	TAG	Gln	Ser	Lys	Glu	Thr	Thr	Ala	Thr	2280
Gln	Ser	Lys	Glu	Thr	Thr	Ala	Ala	Gly	Asp	Ser	***									2340

The 1,173 base pairs in the human IL-16 cDNA. The first Met residue, the five Cys and the GLCF-motif-like repeats are boxed. Dots above corresponding proteins indicate the three nucleotides not present in the previously published sequence⁵. Details of our sequence determination are available on request.

protein (encompassing the DHR3 module of the 40,000- M_r LCF) to CD4, causing some activating events and cytoskeletal reorganization. This could explain the upregulation of interleukin-2 receptors, and the general increase in CD4⁺ cell motility^{2-5,11,12}; antibody blocking of CD4 would diminish these effects. Moreover, blockade of CD4 by LCF/IL-16 binding (even if this were a nonspecific event) could inhibit HIV entry into these cells, perhaps explaining the recent observations of Baier *et al.*¹. Although this does not necessarily detract from the potential therapeutic use of this polypeptide, the notion that it represents an endogenous cytokine⁵ useful as a natural HIV inhibitor¹ must be scrutinized further.

J. Fernando Bazan

Thomas J. Schall

Departments of Molecular Biology and Immunology, DNAX Research Institute, Palo Alto, California 94304, USA

SIR — Cruikshank *et al.*⁵ described interleukin-16 as a 130-residue protein encoded by one main open reading frame of a 2,150-base-pair-long complementary DNA. A biologically active protein is expressed by this DNA even when it has a truncated 5' end (as originally cloned)⁵, indicating that one functional domain of interleukin-16 is in a 390-base-pair coding region. A 130-amino-acid fragment derived from *Escherichia coli* based on this sequence is

active as a lymphocyte chemoattractant⁵ and as an inhibitor of HIV replication¹.

Nevertheless, our nucleotide sequence comparison of interleukin (IL)-16 cDNA clones from non-human primates reveals that the previously postulated ATG start codon⁵ is mutated in two out of six species tested — saimiri and aotus monkeys. Because evolutionarily highly conserved genes probably use the same ATG, we decided to reanalyse the published 5'-untranslated region⁵. We found one large 1,173-bp open reading frame, which includes at its 3' end the 390 bp of the original IL-16 sequence⁵ (see figure).

The first ATG in this open reading frame at position 52 is probably where translation starts¹⁴, and would therefore represent the most likely start codon of this putative IL-16 precursor (prIL-16),

with a predicted relative molecular mass of 39,358. This protein is detectable in correspondingly transfected COS-7 cells with a molecular mass of about 42,000 (M.B. *et al.*, manuscript in preparation).

The prIL-16 sequence contains five cysteine residues and two repeats, which resemble GLGF motifs¹⁵ (see figure). Sequence database searches reveal no significant homologies between the prIL-16 sequence and that of other proteins (except those that carry the GLGF motif). As for the originally published IL-16, there is no consensus signal sequence in prIL-16 (also missing, for example, in IL-1 α and - β) and the secretion pathway for prIL-16 or its processed products is unknown.

Both the 130-amino-acid and the naturally purified versions of IL-16 have a relative molecular mass of 17,000 as measured by SDS gels⁵. This method would not exclude small size differences between the proteins. Although the primary structure of prIL-16 does not resemble a secreted protein, it is conceivable that processing of prIL-16 could lead to the 17,000 form of IL-16 that is detectable in cell culture supernatants of histamine- or mitogen-induced CD8⁺ cells¹⁶. Knowledge of the precise amino terminus of naturally processed and secreted IL-16 is important for investigating its chemoattractant and anti-HIV activities, as altered amino termini can be critical for the biological properties of proteins. Processing of the native prIL-16 may well be essential for generat-

ing properly folded IL-16, which is difficult to obtain in *E. coli* expression systems¹. We are now investigating the processing of prIL-16 and the properties of both prIL-16 and the carboxy-terminal domain of IL-16 from primary lymphocytes to elucidate the biological functions of this protein.

Norbert Bannert

Michael Baier

Albrecht Werner

Reinhard Kurth

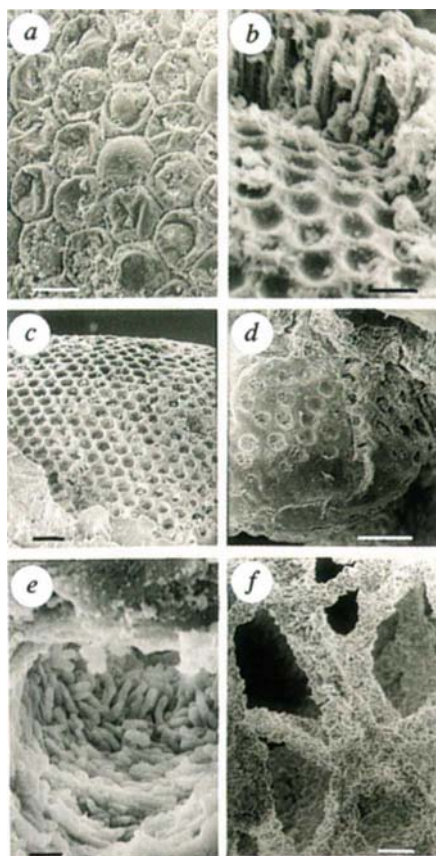
Paul-Ehrlich-Institut, Paul-Ehrlich-Str. 51-59, D-63225 Langen, Germany

Three-dimensionally preserved insects

SIR — The Tertiary limestones of Riversleigh (northwest Queensland) have yielded an assemblage of beetles (including larvae), flies and millipedes preserved in three dimensions, in addition to the already diverse range of vertebrates so far reported¹. These arthropods were recovered from the Upper Site (Late Oligocene/ Early Miocene), which represents a small, shallow, lime-rich pool in a tropical rainforest¹. Scavengers were inhibited by the high salinity, and the conditions for exceptional preservation were enhanced by microbial mats^{1,2}. The arthropods are preserved by mineralization in calcium phosphate (carbonate fluorapatite) without collapse or compaction. They were recovered by acid digestion of the vertebrate-packed limestone; they are preserved in remarkable detail, particularly of the eyes. The only other well-known three-dimensionally phosphatized terrestrial arthropods are from the slightly older Quercy Phosphorites of France, but these specimens have been diagenetically recrystallized or encrusted³.

Replication of soft tissues is more rapid in calcium phosphate than in any other authigenic mineral, and therefore preserves the highest fidelity of detail⁴⁻⁶. The

1. Archer, M., Godthelp, H., Hand, S. J. & Megirian, D. *Aust. Zool.* **25**, 19-65 (1989).
2. Gall, J.-C. *Lethaia* **23**, 21-28 (1990).
3. Handschin, E. *Schweiz. Palaeont. Abh.* **B64**, 1-23 (1944).
4. Briggs, D. E. G., Kear, A. J., Martill, D. M. & Wilby, P. R. *J. geol. Soc. Lond.* **150**, 1035 (1993).
5. Martill, D. M. *Nature* **346**, 171 (1990).
6. Wilby, P. R. *Kaupia* **2**, 99 (1993).
7. Kinzelbach, R. *Zool. Jb. Anat.* **Bd 84**, 539-684 (1967).
8. Clarkson, E. N. K. *Palaeontology* **22**, 1-22 (1979).
9. Paulus, H. F. in *Arthropod Phylogeny* (ed. Gupta, A. P.) 299-383 (Van Nostrand Reinhold, New York, 1979).
10. Clarkson, E. N. K. *Palaeontology* **9**, 464-487 (1966).
11. Horvath, G. & Clarkson, E. N. K. *J. theor. Biol.* **160**, 343-373 (1993).
12. Lucas, J. & Prévôt, L. in *Taphonomy, Releasing the Data Locked in the Fossil Record* (eds Allison, P. A. & Briggs, D. E. G.) 389-409 (Plenum, New York, 1991).
13. Wright, V. P. *Sedimentology* **33**, 831-838 (1986).
14. Jones, B. & Khale, C. F. *J. sedim. Petrol.* **56**, 217-227 (1986).
15. Kellner, A. W. A. *Nature* **379**, 32 (1996).
16. Martill, D. M. & Unwin, D. M. *Nature* **340**, 138 (1989).
17. Wilby, P. R. & Martill, D. M. *Hist. Biol.* **6**, 25-36 (1992).



Scanning electron micrographs of Tertiary beetles from Riversleigh. *a*, Detail of surface of compound eye with intact cornea (QMF34581). *b*, Fractured surface showing rhabdomal walls (top) and exposed rhabdomal 'cups' (bottom) (QMF34582). *c*, Rhabdomal framework underlying the lenses and cones (QMF34583). *d*, 'schizochroal-type' eye of beetle larva (the rhabdomal framework has been lost, exposing the irregularly packed cups) (QMF34584). *e*, Rod-shaped bacteria lining the walls of the rhabdom emplacement in the eye in *c* (QMF34583). *f*, Strands of fungal hyphae criss-crossing the interior of a beetle larva carcass (QMF34584). Specimens gold-coated. Material held by Queensland Museum, Brisbane. 'QMF' denotes catalogue numbers, Queensland Museum. Scale bars: *a*, 10; *b*, 20; *c*, 15; *d*, 40; *e*, 8; *f*, 40 μm .

preservation of insect eyes from specimens at the Riversleigh site is restricted to the Coleoptera. The characteristic hexagonal packing of the compound eye is apparent where the cornea is preserved intact (*a* in the figure). Where the eye is broken the rhabdomal walls may be evident in cross-section and the 'cups' beneath (*b* in the figure). With the lens covering and underlying crystalline cones lost, the rhabdomal framework is exposed (*c* in the figure). The larva of one of the beetles displays an unusual form of compound eye, consisting of large, separate circular lenses, rather than the more usual closely packed hexagonal type. Where the cornea has been lost, it exposes an irregular lattice of 'cups' (*d* in the figure). This 'schizochroal-type' eye occurs only in a small number of living insects⁷⁻⁹, where

the large, separate, circular lenses are thought to maximize light reception. This eye has not been reported previously in a fossil insect, but is characteristic of the extinct marine phacopid trilobites^{10,11}.

Some constraint on bacterial decay is clearly a prerequisite for the fossilization of soft tissues. But it has been shown experimentally that microbial activity is essential to mobilizing phosphate for mineralization¹². The bacteria themselves are mineralized in the Riversleigh insects, in the interior of the rhabdom in the eye, for example, where the pigment cells which lined it have decayed (*e* in the figure). Indeed, the most labile tissues have been largely destroyed; only occasional fragments of muscle are preserved.

Fungi are also preserved in the Riversleigh insects: strands of fungal hyphae can be seen criss-crossing the interior of the head and trunk of the beetle larvae (*f* in the figure). Although fungal hyphae are preserved in some fossil soils^{13,14}, this is the first report of mineralized hyphae

within the carcass of a fossil terrestrial animal.

Reports of mineralized soft tissue in terrestrial animals are rare^{15,16}. The Riversleigh arthropods, with their thicker cuticle, presumably lived (or bred) near the pond, and so would represent only a small fraction of the insect fauna that lived there. Their preservation at the Upper Site may reflect the availability of phosphate in the lake sediment as a result of the accumulation of large quantities of vertebrate material.

Phosphatized arthropods have also been reported in association with vertebrate carcasses in marine settings¹⁷. Wherever there is an accumulation of phosphatized vertebrate material, there will be a high probability that delicate invertebrates can also be recovered.

I. J. Duncan

D. E. G. Briggs

Department of Geology,
University of Bristol,
Queen's Road,
Bristol BS8 1RJ, UK

The nature of the solar wind

SIR—Thomson *et al.*¹ suggest that the wind flowing from the Sun contains many discrete, identifiable modes corresponding to solar internal pressure (p) and gravity (g) modes, in sharp contrast to the standard paradigm of the solar wind as a turbulently evolving magnetofluid with a continuous spectrum of modes. Here we argue for the continuum assumption, and show that the spatio-temporal variation of the solar surface and atmosphere will not allow well-defined modes to persist in the wind, except at low frequencies, where solar-rotation periods and their harmonics account for at least most of the spectral peaks.

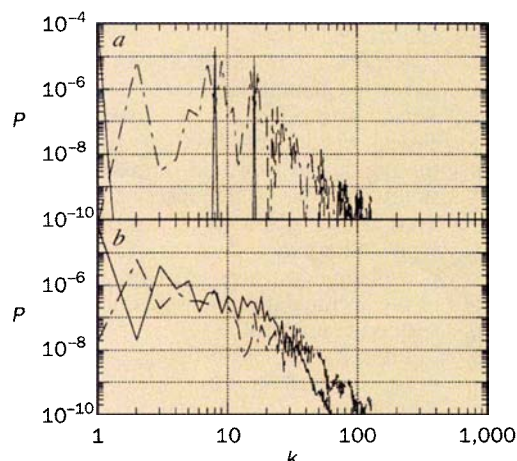
Spectra of magnetic and other solar-wind fields based on the assumption of a continuum of modes frequently reveal power laws, $P \propto f^{-\alpha}$, often with $\alpha \approx -5/3$, characteristic of Kolmogoroff fluid turbulence. Many spectral determinations^{2,3} have shown no higher-frequency preferred peak near the typical 5-min period for p modes, and in general we do not find peaks that are repeatable from one spectrum to the next, or from one component of the magnetic field to the next. Because of the techniques used, these studies are not definitive, so we looked for other evidence.

Solar-wind fluctuations undergo an evolution in which a spectrum that was initially quite flat (down to periods as short as a minute) becomes Kolmogoroff-like⁴. This process also changes the mode of propagation from that of fairly pure Alfvén waves propagating outwards from the Sun to a mixture of modes⁵. Thus, it

is clear that fluctuations at p -mode frequencies have been processed in the inner heliosphere, and if pure modes are to be observed, they must survive this evolution.

Spectra of solar-wind speed and density data from the Ulysses mission (to be presented elsewhere) show harmonics both of the coronal hole rotation period and of a longer rotation period characteristic of the photospheric source region. The peaks are not sharp, but rather have fractional widths of about 5% that can be explained by the nonstationarity and inhomogeneity of the source. Most low-frequency spectral peaks can be explained by these harmonic series or beats between them; this does not rule out g modes, but implies that much more detailed analysis is needed to determine their nature if they are present.

Compressions and rarefactions will broaden the original frequency as observed by the spacecraft, thus eliminating high- Q modes (where Q is the ratio of circulating energy to energy dissipated per cycle). More formally, for a one-dimensional fluid of density ρ and speed $u \gg c$ (the speed of waves in the medium), the frequency of waves in the spacecraft frame will be proportional to u/λ , where λ is the wavelength of the waves. For simplicity, suppose that initially $\rho \propto 1/\lambda$ and thus $f \propto \rho u = \text{const}$. Neglecting pressure and magnetic forces, the initial equation of motion will be $(\partial(\rho u)/\partial t) = -\rho u(\partial u/\partial x)$, and thus, wherever there is a divergence of the velocity, the frequency will be altered



a, Power spectrum of the initial magnetic field (solid) and field magnitude (dot-dash) for a compressible MHD simulation. b, Same as a but at a later time.

from its original value. (For waves compressed with the plasma, $l \propto 1/\rho$ is maintained.)

The figure (a) shows the initial spectrum of a magnetic component of three small-amplitude magnetohydrodynamic (MHD) waves and the corresponding spectrum of the field strength for a one-dimensional MHD simulation to test these ideas. The initial field magnitude spectrum, important in the modulation of energetic particles, is different from the component spectrum, both in that peaks are modulated and thus shifted and in that it is more continuous. The figure (b) shows the spectrum after the equivalent of about a day of evolution in the solar wind for modes at p -mode frequencies; no higher-wavenumber features remain simply due to steepening and weak nonlinear interaction. The equality of optical and spacecraft higher-frequency p -mode observations¹ still poses a problem, but we believe that the case for random coincidences should be re-examined in light of the intermittent nature of the solar-wind turbulence.

Finally, there is no known mechanism for the required g - and p -mode coupling to any quantity measured by spacecraft. This, in addition to the considerations above, means that the arguments are very strong for a continuum of solar-wind modes with the only preferred basic frequencies corresponding to solar rotation.

D. Aaron Roberts
K. W. Ogilvie
M. L. Goldstein

*Laboratory for Extraterrestrial Physics,
NASA Goddard Space Flight Center,
Greenbelt, Maryland 20771, USA*

THOMSON ET AL. REPLY — We appreciate the remarks by Roberts *et al.*, but disagree with their conclusions. Roberts *et al.* state that the standard paradigm of the solar wind is a turbulently evolving magneto-fluid with the consequence that the solar

wind has a continuous spectrum. Indeed, time-series variations of the interplanetary magnetic field and of galactic cosmic rays have been studied to gain an understanding of phenomena such as the turbulent propagation of the charged particles⁶ and the hypothesized fractal nature of the field configuration⁷. We argue that this ‘continuum hypothesis’ has never been subjected to a careful statistical test against an alternative theory suggesting that discrete modes were present in the solar wind, because such a possibility had never been considered.

In our paper¹, we stated explicitly “We take as our null hypothesis, H_0 , the standard assumption that the various time-series considered have a purely continuous spectrum with no discrete components apart from solar rotation.” All the tests we reported, as well as many others, had extremely unlikely outcomes under H_0 .

Given that spectra of time-series of interplanetary data are commonly computed, why have discrete frequencies not been identified before? The reason is that those making the measurements were expecting continuous spectra, so chose estimation methods appropriate for determining parameters of a continuous spectrum, such as the slope. Most of these estimation methods used a single data window, so only large-amplitude periodic terms would have been detectable. It has frequently been observed that such spectra often have larger peaks than would be reasonably expected from a process with a continuous spectrum. But because the locations of these peaks tended to change between samples, it was believed that they were not the result of an embedded periodic process.

In retrospect, the “unreasonable” statistical behaviour of such spectra is completely understandable in terms of our hypothesis. The spectrum of solar modes is quite dense in both the p - and g -mode ranges, so the resolution requirements are severe: of the 2,241 theoretical g -mode frequencies for $l = 1$ to $l = 6$ inclusive (D. Gough, personal communication), there are 1,071 adjacent frequencies where the beat period exceeds 1 year. A spectrum

estimated from a record of a few months’ duration will almost certainly contain some frequencies where two (or more) of these unresolved modes are ‘in phase’ and so will have peaks at those frequencies. At a different time, these modes may be out of phase, while those at some other frequency may be aligned. Thus, such spectra will seem unreasonably variable compared with the variability expected under H_0 .

Previous studies used inefficient statistical methods for estimating spectra. For example, of the two studies cited by Roberts *et al.* as not showing “any higher frequency preferred peak[s] near... 5 min period”, their ref. 2 used a Blackman and Tukey method, and their ref. 3 used an unwrapped fast Fourier transform. Both methods are badly biased⁸ and, in our view, should be avoided⁹. Nonetheless, several of the spectra for solar-wind data acquired at 0.3, 1.0 and 10 AU of ref. 3 of Roberts *et al.* have clear peaks in the millihertz range as well as in the ‘intermediate range’ of frequencies of Table 4 in our paper¹. There are many other papers, for example refs 2, 10, in which peaks in the millihertz region are reported.

In support of our hypothesis that solar modes are responsible for these discrete lines, there is very close agreement between the frequencies reported in Table 2 of our paper and theoretical predictions of g -mode frequencies¹¹, as well as some optical measurements^{12,13} of possible g modes.

To conclude, we believe that it is an observational fact that the solar wind includes many discrete modes. We do not dispute the fact that turbulence effects also occur; as we noted¹, the ions have become turbulent at p -mode frequencies at 5 AU, while the electrons have not. The Voyager data show that some discrete frequencies are still identifiable at 10 AU but they appear to be mixed by 20 AU (ref. 14). The analysis of such discrete modes should aid in developing better theories of the solar wind.

David J. Thomson
Carol G. MacLennan
Louis J. Lanzerotti
*Bell Laboratories,
Lucent Technologies,
Murray Hill,
New Jersey 07974, USA*

1. Thomson, D. J., MacLennan, C. G. & Lanzerotti, L. J. *Nature* **376**, 139–144 (1995).
2. Matthaeus, W. H. & Goldstein, M. L. *J. geophys. Res.* **87**, 6011–6028 (1982).
3. Goldstein, M. L., Roberts, D. A. & Fitch, C. A. *J. geophys. Res.* **99**, 11519–11538 (1994).
4. Bavassano, B., Dobrowolny, M., Mariani, F. & Ness, N. F. *J. geophys. Res.* **87**, 3617–3622 (1982).
5. Roberts, D. A., Goldstein, M. L., Klein, L. W. & Matthaeus, W. H. *J. geophys. Res.* **92**, 12023–12035 (1987).
6. Owens, A. J. & Jokipii, J. R. *J. geophys. Res.* **79**, 895–906 (1974).
7. Burlaga, L. F. *J. geophys. Res.* **97**, 4283–4293 (1992).

8. Thomson, D. J. *et al. Phys. Earth planet. Inter.* **12**, 217–231 (1976).
9. Thomson, D. J. *Phil. Trans. R. Soc. A* **330**, 601–616 (1990).
10. Goldstein, M. L., Burlaga, L. F. & Matthaeus, W. H. *J. geophys. Res.* **89**, 3747–3761 (1984).
11. Guenther, D. B., Demarque, P., Kim, Y.-C. & Pinsonneault, M. H. *Astrophys. J.* **387**, 372–393 (1992).
12. Hill, H. A. & Yi-min, G. U. *Science in China A* **33**, 854–866 (1990).
13. Frohlich, C. *et al. The Sun in Time* (eds Sonett, C. P., Giampapa, M. S. & Matthews, M. S.) 11–29 (Univ. Arizona Press, Tucson, 1991).
14. Lanzerotti, L. J. *et al. Astrophys. J.* **380**, L93–L96 (1991).

Science struck dumb

N. J. Mackintosh

The g Factor: General Intelligence and Its Implications. By Christopher Brand. Pp. 253. To have been published by Wiley*.

HERE is a novel way for a publisher to get free publicity for a newly published book: immediately withdraw it from publication, because you find some of the author's assertions repellent and have no wish to support his views by disseminating them.

I quote freely from a press release issued on 17 April by the president and chief executive of John Wiley, dissociating his company from Christopher Brand's *The g Factor*, due to be published the following day. This seems a singularly cack-handed attempt at censorship, and although the managers of Wiley may be congratulating themselves on the purity of their collective conscience, others may question their good sense, competence and integrity. How is it that they found out about the repellent nature of Brand's views only *after* they had printed and distributed copies of the book in which he expressed them? Do they not appreciate that attempts at suppression or censorship usually only enhance the reputation of otherwise rather undistinguished books (think of *Lady Chatterley's Lover*)?

Spirited defence

So what are Brand's abhorrent views? This is, of course, a book about IQ tests, and it will come as no surprise to learn that Brand thinks that IQ tests are a pretty good thing. The book is largely a spirited defence of four propositions. First, that IQ tests really do measure intelligence, and that the critics' arguments to the contrary are refuted by the claim that all attempts to find other measures of intelligence uncorrelated with IQ scores have failed. Second, that intelligence is a unitary process, largely reducible to speed of apprehension or intake of simple perceptual information. Third, that genetic differences are much more important than environmental differences as a cause of differences in IQ — both between individuals and between certain groups. And finally, that all this matters, because the intelligence measured by IQ tests is not just 'academic' intelligence, but something of much wider practical significance.

The defence of these propositions is liberally interspersed with even more spirited attacks on a variety of bogeymen, which include not only such expected

targets as Stephen Jay Gould and Steven Rose, the massed battalions of the politically correct, and those who have sought to deny the possibility of estimating the heritability of IQ by insisting that genetic and environmental influences are inextricably intertwined and interact in ways too complex to understand, but also most other branches of psychology, including behaviourism, cognitive psychology and other less familiar brands labelled idealism and constructivism. And Brand is refreshingly willing to startle or shock his readers, as when he tells us of a particular idealistic psychologist from San Diego, with a sustained, scholarly interest in helping children with learning difficulties, whose great love is classical music and whose hero is Mahatma Gandhi. Who is this paragon of virtue? None other than Arthur Jensen.

Although this second aspect of the book will no doubt annoy some readers, there is little harm in that. It is the positions Brand defends that will raise more legitimate concern. There is, no doubt, much to be said for stating a position forthrightly and unambiguously. But that is not the same as going beyond the evidence. Not many readers will accept that the evidence justifies some of Brand's assertions. With little or no hesitation, he settles on 0.75 as an estimate of the broad heritability of IQ: this is at the upper limit of the plausible, and for some time now most behavioural geneticists have been content with a very rough estimate of 0.50. Even more startling, if less likely to offend political sensibilities, is his confident assertion that IQ scores correlate 0.75 with performance on 'inspection time' tasks. This is the basis for Brand's claim that the nature of g, or general intelligence, has now been identified as speed of apprehension. I know of few other investigators who would claim that the correlation is greater than 0.50, and my own reading of the evidence suggests that a figure of 0.35–0.45 is much nearer the mark. Nor has Brand, or anyone else, ever established that inspection time owes its correlation with IQ scores to its correlation with g.

If you want to offend political sensibilities, there is no better way than to assert that in the United States blacks obtain lower average IQ scores than whites, and that this difference cannot be attributed to environmental differences. Although

many would like to draw a discreet veil over this topic, there is no sensible way to deny the truth of the first of these propositions. But why should we believe that the difference is probably genetic in origin? Brand seeks to persuade the reader of this, and, I suspect, has succeeded in persuading himself, by the tactic of rubbishing so many other arguments advanced by the critics of IQ tests that one ends up persuaded that here too the politically correct critic has simply got it wrong. The tactic may be persuasive, but it should be resisted.

Wishful thinking

It is true that much criticism of IQ testing has been ignorant, tendentious and based on wishful thinking; contrary to the critics' arguments, IQ does have significant heritability, and IQ scores do predict, rather better than parental socioeconomic status, some reasonably important aspects of people's achievements. But it really does not follow that we can also dismiss the claim that black-white differences in average IQ may well be entirely environmental in origin. A moment's reflection should surely suggest how extraordinarily difficult it will be to rule out this possibility. The standard way to ascertain the role of genetic factors is to hold environmental differences constant. So to show that black-white differences in IQ are genetically caused, we should need to bring up black and white children in comparable environments. And how, in a racist society, are we going to do that?

Some people, perhaps including the New York managers of Wiley, will regard Brand as little better than a reactionary, racist ideologue. But that would not do justice to many of his views. In his last and longest chapter, he develops a strong argument for individual freedom of choice in education. Pointing out that the top 10 per cent of seven-and-a-half-year-old children obtain higher IQ test scores than the bottom 10 per cent of fifteen-and-a-half-year-olds, he asks why we persist in segregating schoolchildren by chronological age rather than by mental age. Warming to his radical libertarian theme, he suggests that we should simply allow children to choose the intellectual level of their classes. Because it turns out that adults choose their friends, and spouses, by IQ (without of course knowing), he confidently predicts that schoolchildren would do so too. It is an idea at least worth injecting into debates on education. □

N. J. Mackintosh is in the Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, UK.

*For further details see this week's News section.

Cyborg discourse

William Clocksin

The Closed World: Computers and the Politics of Discourse in Cold War America. By Paul N. Edwards. MIT Press: 1996. Pp. 440. \$40, £33.95.

How did global military control come to seem a practical technological proposition? How did tradition-bound military services, with their models of leadership based on battlefield experience, become transformed into managers of automated systems implementing preprogrammed plans based on abstract strategies? What held the official US strategic discourse of the Vietnam era together, in the face of what could have been construed as glaring evidence of failure? Edwards tackles questions like these arising from the tangled web of politics, culture and computer technology. Technology, tools and culture cannot be understood in isolation; just as tools such as weapons and computers have emerged from the culture, so too have tools caused profound changes in the culture. In particular, computers and the political imagination have reciprocally extended and transformed each other in ways that are responsible for how we see ourselves as humans.

Paul Edwards begins by describing the emergence of a "closed-world discourse" of global surveillance and control through high-technology military power. Closed-world discourse reduces any problem to a system having a known specification, rationally ordered to produced carefully defined outputs. Everything in the closed world becomes a system, composed of subsystems and integrated into supersystems. These nested systems are constituted in and through metaphors, technologies and practices. The metaphors are information, communication and programming; the technologies are computation and control; and the practices are abstraction, simulation and panoptic management. Computers helped to create and sustain this discourse by allowing the practical construction of central real-time military command and control systems on a vast scale, and by helping the metaphorical understanding of world politics as a sort of system subject to technological management.

Closed-world politics saw the West as a world enclosed within its defences, the Soviet Union as a closed world to be penetrated, and the globe as a world enclosed within the capitalist-communist struggle. Each of these versions of the closed world found an embodiment in computerized

command and control systems. The Cold War political goal of 'containment' led to, for example, the SAGE continental air defence system, the Strategic Air Command Control System, and the World-Wide Military Command and Control System. Edwards tells the stories of how and why these systems were developed and the interrelationships between military research needs and the origin of cognitive science and artificial intelligence research. The reductionist 'closed world' view of cybernetics provided powerful images not only for political use, but also for individualistic, mechanistic, cognitive models of human behaviour, and for society at large.



Military might — the cover of *Time* magazine, 23 January 1950, speculating about thinking machines.

For example, exhibits of life in the future displayed at the 1964 World's Fair represented closed artificial systems in hostile environments, such as space capsules, undersea communities inhabiting pressurized domes, underground houses, and space stations. The displays depicted smiling prosperous white families going about their daily activities inside these bubble-like shells. Such simplistic images of personal security through containment and control implicitly neglected the values of coming to terms with the complexities of Earth together with all its inhabitants and their needs, and reached an apotheosis in the Strategic Defense Initiative of the Reagan era.

By the late 1950s, cognitive psychology and artificial intelligence seemed increasingly to be parts of a single whole, united through the abstraction of symbolic processing, and giving rise to what Edwards

calls 'cyborg discourse'. For the cyborg — from 'cybernetic organism' — components of an intelligent system, whether man or machine, are considered to be interchangeable. By understanding humans and computers as information-processing machines, the reductionism of cyborg discourse helped to integrate people into the complex technological systems required for military purposes in turn fostered by closed-world ideology. Yet, far from seeing this as merely a research problem concerning human-computer interfaces, Edwards shows through an analysis of popular films such as *Star Wars* and *Terminator* how much of social life and our perceptions of human identity (what we think we are) has become invested with cyborg discourse.

Edwards's treatment of cyborg identity is informed by the postmodern condition: the radical subjectivity of the closed world permits neither escape nor transcendence, and offers only the possibility of reconstitution. He proposes that the sole means of genuine self-determination is the "political subject position of the cyborg", which can find "a habitation, if not a home, inside the closed world". Yet Edwards seems not to handle the paradox that cyborg discourse is itself a manifestation of late modernism, the last microchip-powered gasp of the failed enlightenment and romantic projects. Cyborg discourse is conditioned by precritical endorsement of the kind of reductionistic mind-body dualism that hasn't changed substantially since even before Mary Wollstonecraft Shelley's *Frankenstein*. Cyborg discourse is at least as vulnerable to deconstruction (the entirely legitimate project of exposing and challenging tacit assumptions and taken-for-granted realities) as Edwards shows that Cold War discourse is.

There is a little something in this book for almost everybody. Military historians will find some analysis of source material relating to the influence of computer technology on the formation of doctrine, political policy and interservice relations.

Cognitive scientists and artificial intelligence researchers will learn more personal details and anecdotes about those who contributed to the origins of their subjects. In this connection, although Edwards shows convincingly that computers are best understood in the context of their roles in providing metaphors and political icons, he has probably been ill-advised on practical matters of computer science: however far a point can be stretched, John McCarthy did not "invent" time-sharing, and expert systems do not date from the mid-1960s.

Cultural critics will find an engagement with the notions of the perpetual conflict of

the closed world versus the unbounded natural setting of the 'green world' (although to be fair the closed-green contrast does not play a central role in the book), and an introduction to popular film as a vehicle for cyborg discourse. There are passing nods in the direction of Foucault, Derrida, social constructionism and narrative hermeneutics, but the relevant issues are not handled in any depth. The epilogue nicely shows how closed-world politics was reasserted in the Gulf War of 1990–91 to become "an emblem for the glories of America's waning global hegemony", but the following discussion of the World Wide Web offers little more than journalistic cliché: "the new connectionist cyborgs will have to pay their way at the toll booths of the information superhighway".

Edwards's main contribution is to illuminate the subjective, contingent, interconnected structure of science, politics and military technology in Cold War America. Edwards makes a convincing case that readers should seriously question whatever illusions they might have about the practicality of automatic global control, about the ethical neutrality of science and technology, and about why these practices are funded. But, like cyborgs, illusions die hard, and one of the messages of this book is that we humans are inclined to replace old illusions with new ones. □

William Clocksin is in the Computer Laboratory, University of Cambridge, Cambridge CB2 3QG, UK.

Alternative histories

S. Nomanul Haq

Science and the Raj, 1857–1905. By Deepak Kumar. Oxford University Press: 1995. Pp. 273. £15.99.

HISTORIANS of science have sometimes spoken of Aristotle's obsession with taxonomy and classification. They explain it by observing that this Greek 'giant' remained essentially a biologist, for even in his logic and metaphysics he maintains a classificatory approach. Read this animated writing from Deepak Kumar, and you will begin to appreciate the philosophical cogency of Aristotle's obsession, for *Science and the Raj* immediately throws into sharp relief a fundamental question of classification: does this book constitute history of science or part of the history of British colonialism?

If one has to evaluate the worth of this book this is not a rhetorical question; for if *Science and the Raj* is concerned with the imperialist machinations of the Raj, it is without doubt a valuable addition to the genre of research to which it belongs. The book is valuable despite its somewhat triv-

ial thesis, and the reader will find a rich treasure of tightly packed first-hand information derived from primary sources. But as a study in the history of science it is at best naive; at worst it is misleading and sometimes simply wrong; the historian of science should be prepared to be disappointed.

Looking through the window of the history of British colonialism in India, one would truly admire Kumar for all the variegated sights he brings before us: the tight government controls on scientific research; the nakedly profit-oriented approach of the colonial scientists; the stress on technology rather than on theoretical work; the insensitive exploitation of

The British bacteriologist Sir Ronald Ross, who discovered the transmission of malaria by mosquitoes, on the steps of his laboratory in Calcutta, 1898, with his wife and assistants. Cages for malarial birds appear in the foreground.

natural resources; the botanical and geological surveys for economic and political gains; the mockery of the local scientific traditions; the frustrations expressed at the "inherent limitations of the native"; the suppression of local talent and promise. All these constitute an exposé of colonial attitudes in India. What is more, the historical evidence here is solid, plentiful and primary; with sustained vigour and narrative clarity, episodes pass before us and in the process form a pattern — a pattern which Edward Said, whom Kumar cites, has called 'Orientalism': both a mode of action and a mode of thought.

Considered as a work in colonial history, Kumar has an unmistakable thesis but it is one that is much too familiar: that the colonial officials were on the whole exploitative. He tends to qualify this position but it is unfortunate that he hardly develops his qualification, nor does he speak lucidly on this matter. For example, at the very outset he raises an important question: "[T]he common sense of an alien power could scarcely have thought of anything other than profit. But can this be the final verdict?" Here is his answer: "The more one studies the archival and contemporary sources, the more one finds various shades surfacing at different levels

of the official hierarchy as well as non-official forums". But there is no coherent development of this rather verbose answer. On the other hand, the book contains some powerful, though scattered, insights. For example: "Guns and sails could win territories but the [colonial] empire could be sustained only by emphasizing differences and operating from a... pedestal"; again his observation is highly significant: "The cultural projects [of colonialism] showed a deeper penetration and a greater resilience than its economic forays".

But it is ironic that Kumar himself fell victim to these very "cultural projects". And here I begin to look at his book from

Wellcome Institute Library

the other window, that of the history of science. It is obvious that the author's view of science is what his Victorian masters taught during their Raj — a view that is positivistic; a view in which the ghost of Francis Bacon's naive inductivism looms large; a view that has long been discarded in London. To begin with, Kumar is unable to distinguish between science and technology, and admits this handicap he therefore confuses technological and industrial development with scientific growth. And science for him is solely what survives in today's modern science, something that has

"universality" and "utility". So any attempt on the part of the "natives" to reclaim their scientific heritage is for him "revivalism". It is for this reason that he fails to appreciate the tremendous scientific worth of the work of the fifth-century Indian astronomer Āryabhata.

As for the question of utility, Kumar ought to understand that advances in science are theoretical advances, without regard to utility or application. The case of Copernicus, cited at several places in the book, is instructive. From the point of view of the calculations of planetary orbits, the Ptolemaic system was in many instances more accurate than the Copernican and so had more utility, and yet it was discarded. Copernicus did not achieve more utilitarian success; rather, he took a daring philosophical step, and brought about a revolution in science. Again, one may legitimately say that for much of our everyday engineering purposes, classical mechanics has more utility than relativistic mechanics, for here the latter effectively reduces to the former. But despite its limited utility, Einstein's system is a mammoth advance in science — a mammoth theoretical advance. Kumar's inability to distinguish between science and technology leads him to misquote his

sources: "Most of the leading scientists of eighteenth-century Europe", he says in citing D. Cardwell, "were instrument-makers by profession". But Cardwell had said, and only in parentheses, that "Some of the leading scientists of the century were instrument-makers by profession" (emphases mine). Was Joseph Priestley an instrument-maker? What about John Dalton and Joseph Black?

To be sure, Kumar the historian of science has no access to primary sources at all; contrast it with Kumar the historian of colonialism. It is embarrassing to express firm views on, for example, the history of Indian astronomy without looking at a single Sanskrit text, not to mention Arabic and Persian writings. And even these secondary sources here are highly unreliable. The case of the astronomer-ruler Maharaja Jai Singh is one glaring instance. Kumar says, on the authority of a poor secondary reference, that the Maharaja had accepted the Copernican model; this is simply wrong. Kumar would be well advised to consult the researches of historians such as K. P. Shukla, V. N. Sharma, S. M. R. Ansari and David Pingree. To measure the worth of the Indian scientific tradition in terms of its proximity to modern science is to betray an unhappy lack of understanding of the scientific spirit and scientific enterprise.

The author also raises the vexed question of science and ideology, and contradicts himself. On the one hand, he quotes with approval the views of al-Afghānī and Sir Syed Ahmed Khan (there is no 'European' science and 'Islamic' science, science is science; science is secular, universal, ideology-free, and so on); on the other hand, he himself talks about "colonial science" — in fact he goes as far as to say: "Colonial science... had an ideology"; in other words, science is not utterly ideology-free. But if this is the case, why does he disapprove of the ideological concerns of the nineteenth-century Indian scientists such as P. N. Bose, P. C. Ray and R. Trivedi who spoke of Hindu culture, civilization, cosmology and metaphysics? Why do they betray "streaks of revivalism"? It seems to me that these people understood the nature of science better than Sir Syed did; and they had a more powerful intellect than their contemporary Chattopadhyaya who declared so naively — and with such a ringing apologetic tone — that the system of Āryabhata is identical to that of Copernicus; this is absurd.

If *Science and the Raj* is about colonialism it is welcome; if it makes any claims to being a history of science, it ought to be ignored. □

S. Nomanul Haq is in the Department of Religious Studies, Brown University, Providence, Rhode Island 02912, USA.

Interrogation and interpretation

Donald J. DePaolo

Introduction to Geochemical Modeling.

By Francis Albarède. Cambridge University Press: 1995. Pp. 543. £80, \$100 (hbk); £29.95, \$39.95 (pbk).

THE focus of most geochemistry is the interrogation of nature with ever more penetrating measurements — higher precision, better spatial resolution, greater sensitivity. The Holy Grail of geochemistry could be said to be the 'triumphant measurement', the implications of which are both profound and patently obvious, if not quite model-independent. Who is to deny that the Solar System is 4.55 billion years old, a statement so elegantly made by lead-isotope measurements?

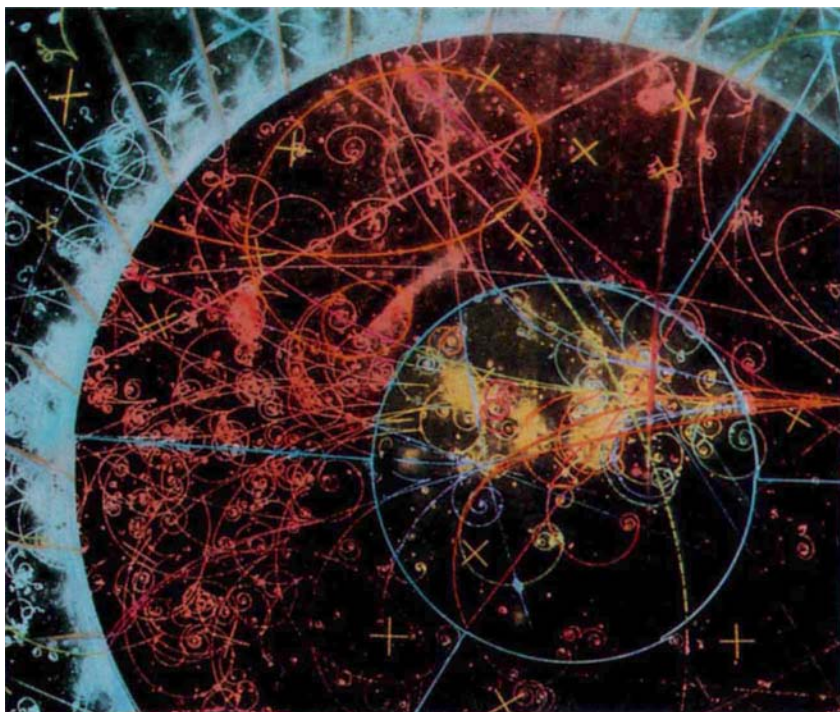
But most geochemical data, especially those that prove to be indispensable for environmental and geological problems, require involved interpretation regardless of the sophistication of the measurements. The ultimate value of geochemical data, no matter how beautiful they appear to the creator and the connoisseur, is no

more and no less than as a constraint on a model, and the poorer the model, the less valuable the data.

Francis Albarède has produced an impressive and much needed contribution to the continuing transformation of geochemistry into an explicit model-based science. Most of the essential mathematics of geochemistry is presented in a clear and readable fashion. The special value of the book lies in its relative completeness, in both the breadth of topics and approaches covered and the step-by-step working of relevant problems.

This is an indispensable handbook for the research geochemist, and could serve as a text for a graduate or advanced undergraduate geochemistry course. Graduate students will find the book invaluable for access to the geochemical literature, as well as for dealing with their own data. Albarède presents mathematics admirably by illustrating the applications to common geochemical problems while revealing his personal vision of quantitative geochemistry and providing a useful introduction to the attendant terminology. □

Donald J. DePaolo is in the Department of Geology and Geophysics, University of California, Berkeley, California 94720-4767, USA.



COLLISION between a high-energy neutrino and the nucleus of a hydrogen atom. Taken from the cover of the paperback edition of *From Quarks to the Cosmos: Tools of Discovery* by Leon M. Lederman and David N. Schramm. The authors cover exciting new developments in the seven years since the book was first published such as the discovery of the top quark in 1995 and the results of the Cosmic Background Explorer satellite in the early 1990s. They also lament the demise of the Superconducting Supercollider project, but remain optimistic about the completion of CERN's proposed Large Hadron Collider. W. H. Freeman/Scientific American Library, \$19.95, £14.95.

The role of gravitational potential energy in active deformation in the southwestern United States

Craig H. Jones*, Jeffrey R. Unruh† & Leslie J. Sonder‡

* Cooperative Institute for Research in the Environmental Sciences, Campus Box 216, University of Colorado at Boulder, Boulder, Colorado 80309-0216, USA

† William Lettis & Associates, Inc., 1777 Botelho Drive, Suite 262 Walnut Creek, California 94596, USA

‡ Department of Earth Sciences, 6105 Fairchild Science Center, Dartmouth College, Hanover, New Hampshire 03755, USA

In orogenic belts, intrinsic (buoyancy) and extrinsic (boundary) forces act on the continental lithosphere; determining the relative importance of these forces is fundamental to understanding the dynamics of continental deformation. In the western United States, estimates of gravitational potential energy show that, given available constraints on rheology, buoyancy forces are sufficient to produce the wide variety of rates and styles of observed deformation, including absence of deformation. This underscores the importance of gravitational potential energy for deformation in orogenic belts.

THE 800–1,600-km-wide Cordilleran orogen of the western United States represents a collage of active tectonic styles¹, including strike-slip motion along the Pacific–North America plate margin in California, extensional deformation in the Basin and Range, contractional tectonics in the Pacific Northwest and west-central California, and uplifted but relatively undeformed crust in the Sierra Nevada and Colorado Plateau (Fig. 1a). Throughout much of the interior, the Cordillera is a broad, elevated plateau with an average surface elevation of 1.5–2.2 km (ref. 2), with some individual mountain ranges attaining elevations of 3.0–4.0 km.

One class of explanations for active deformation of the Cordillera appeals to interactions along plate margins or at the base of the lithosphere. Such explanations include back-arc spreading behind the Cascadia subduction zone³, shear tractions at the base of the lithosphere from mantle convection or subduction processes^{4–6}, distributed northwest dextral shear^{7,8}, and a misfit between Pacific–North American motion and the trajectory of the San Andreas fault system⁹.

A second class of explanations holds that Cordilleran deformation is driven by buoyancy arising from spatial differences in gravitational potential energy (PE)^{10–12}. Such forces are intrinsic to the lithosphere, not externally imposed, and have been suggested as the cause of late orogenic deformation at numerous locations globally^{13–16}. For example, Sonder *et al.*¹¹ showed quantitatively that the gravitational PE stored in thickened continental lithosphere following the end of the Late Cretaceous to Early Tertiary Laramide orogeny was sufficient to cause the extension observed in the Great Basin, provided that the lithosphere was warm enough, and hence weak enough, to extend. To date, however, there has been no examination of the present-day correspondence, or lack thereof, between the distribution of PE and active deformation for the Cordillera as a whole.

We have attempted such an examination by calculating the distribution of lithospheric PE in the southwestern United States, using crustal density structures (derived from seismic refraction profiles) and the assumption of local isostasy. The regional variations in PE correlate well with modern-day deformation regimes, generally predicting the styles and rates of deformation that are observed. Although other forces may be present, the buoyancy forces arising from PE variations are capable of pro-

ducing virtually all active deformation other than the strike-slip motion at the Pacific–North American plate boundary.

Gravitational PE and deviatoric stress

Crucial to an assessment of the forces driving deformation in the Cordillera is understanding (1) how potential-energy contrasts are related to the density distribution in the lithosphere, and (2) how much potential energy is available to drive deformation.

The potential energy per unit area of a lithospheric column is given by the integral of the product of weight (per unit area) and height z above a reference level, usually taken as the level of isostatic compensation. For a lithospheric column of density $\rho(z)$, the potential energy PE is

$$PE = \int_0^{e+L} \rho(z')gz'dz' = \int_{-e}^L \rho(z)gLdz - \int_{-e}^L \rho(z)gzdz \quad (1)$$

where z' increases upwards from an origin at the compensation level; z increases downwards from an origin at sea level, g is gravitational acceleration, L is the compensation depth, and e is the topographic elevation^{13,17,18}. Under the assumption of local isostatic compensation, it can be shown that the difference in PE (ΔPE) between any two columns of continental lithosphere is equal to the difference in the vertical integral of the vertical normal stress σ_{zz} (refs 13, 17):

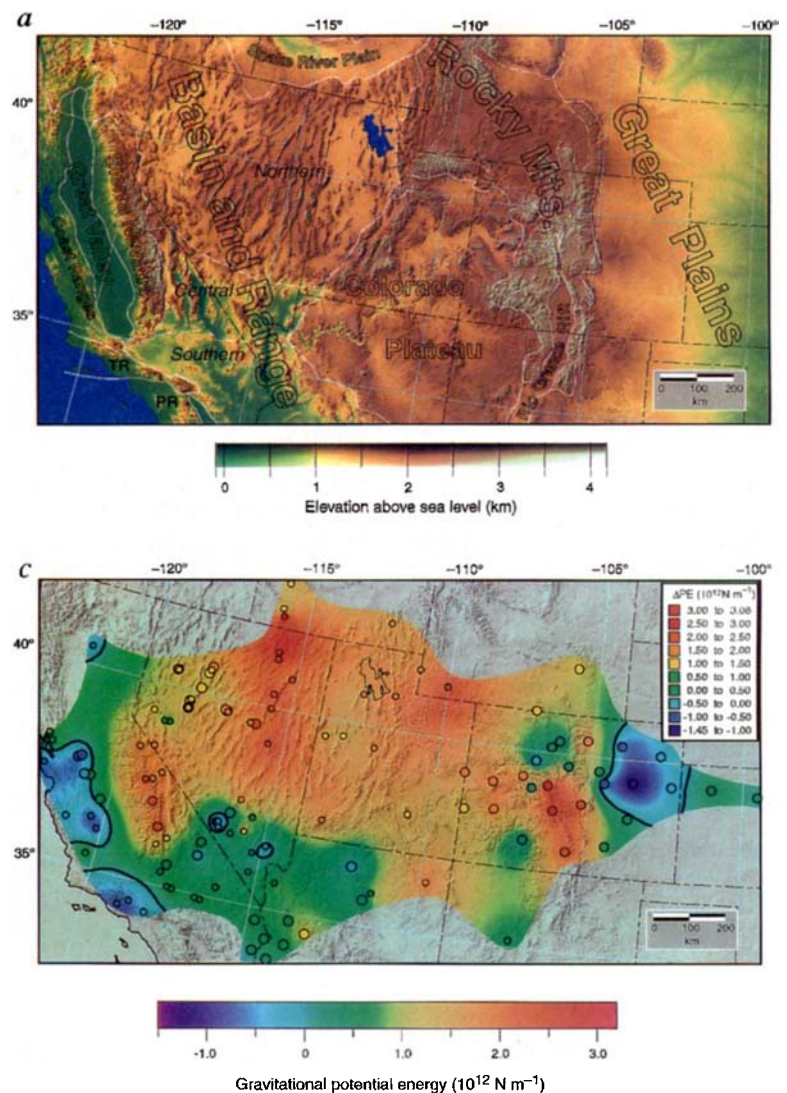
$$\Delta PE = \Delta \left[\int_{-e}^L \sigma_{zz}(z)dz \right] = \Delta \left[\int_{-e}^L \left(\int_{-e}^z \rho(\zeta)g d\zeta \right) dz \right] \quad (2)$$

By vertically integrating the equations of motion, the horizontal equations can be written as

$$\frac{\partial \bar{\tau}_{ij}}{\partial x_j} + \frac{\partial \bar{\tau}_{zz}}{\partial x_i} = \frac{1}{L_0} \frac{\partial (PE)}{\partial x_i} \quad (3)$$

where $\bar{\tau}$ indicates a vertically averaged deviatoric stress, L_0 is the thickness of the lithosphere, and the subscripts i and j refer only to horizontal coordinates. Thus, stresses are directly related to PE. In two dimensions (one horizontal and one vertical) the relationship is even simpler; deviatoric stresses (and strain rates if the effective viscosity is constant) are linearly proportional to differences in PE (ref. 19). The relationship between ΔPE and strain rate is of course more complicated in three dimensions, in the presence of vertical variations in strain rate, or when the rheology

FIG. 1 *a*, Shaded topographic relief map of the western United States with physiographic provinces. (TR, Transverse Ranges; PR, Peninsular Ranges.) Illumination is from the west. *b*, Contribution of the mantle to elevations in the southwestern United States (mantle buoyancy H_m). Dots indicate the locations of seismic profiles used to constrain crustal density columns; dot sizes reflect relative confidence in techniques used in the seismic interpretations (see Supplementary Information). Coloured images in *b* and *c* represent interpolations between data points shown using unweighted linear kriging⁶². *c*, Calculated gravitational potential energy (ΔPE) distribution in the western United States, relative to an asthenospheric reference column, displayed over a shaded relief map. Contour shown separates areas with PE greater than the reference asthenospheric column from areas with PE less than the reference column. *d*, Geoid filtered above degree 7 (tapered below degree 11). The top scale bar indicates the theoretical relation between geoid and ΔPE ; colours used match the scale in *c*.



is not constant. Nevertheless, the ΔPE is a useful first-order measure of the tendency of continental lithosphere to deform owing to internal body forces.

Following Lachenbruch and Morgan²⁰ and Houseman and England¹⁹, we adopt as a reference state for PE an asthenospheric column of density ρ_a with its top at $H_0 = 2.4 \text{ km}$ below sea level and no lithosphere or water above it. A continental column in isostatic balance with the reference column but with greater PE (positive ΔPE) should be in a state of horizontal deviatoric tension, and a column with negative ΔPE should be in compression.

Southwestern US potential energy distribution

Where seismic velocity profiles are available (see Supplementary Information), we assess the crustal contribution to PE by first converting seismic velocities to densities with an empirical relation corrected for a mean geotherm and pressure²¹ for P-wave velocities greater than 6.0 km s^{-1} and a simpler empirical relation²² for velocities less than 6.0 km s^{-1} . We then use equation (1) to directly calculate the crustal contribution to PE relative to an asthenospheric column,

$$\Delta PE_c = \int_{H_0}^{z_c} \rho_a g z dz - \int_{-l_m}^{z_c} \rho(z) g z dz \quad (4)$$

where z_c is the depth of the Moho below sea level. The elevation ε_m averages topography around the measurement point using the

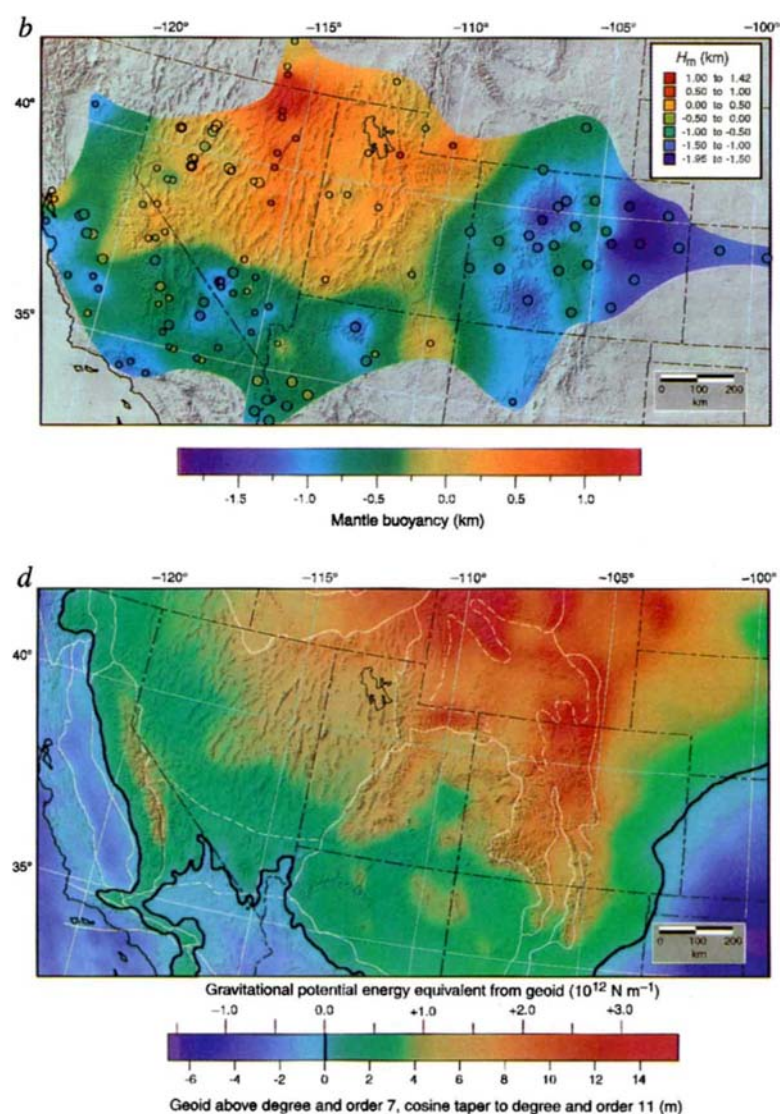
function describing the flexure of a plate from a point load²³

$$\varepsilon_m = \frac{\iint \varepsilon(r, \theta) \text{kei}(r/\alpha) dA}{\iint \text{kei}(r/\alpha) dA} \quad (5)$$

where α is the flexural parameter from ref. 24 and the integration is over a disk of radius $\sim 3.9\alpha$, the distance of the first zero crossing of the kei (Thomson) function.

Ideally, we would prefer also to calculate the mantle contribution of the PE by estimating $\rho_m(z)$ from seismic observations, but seismic velocity estimates in the mantle are poorly constrained. In addition, velocity–density correlations are much poorer for ultramafic rocks in the mantle than for rocks of crustal composition²⁵, a point underscored by the indistinguishable P-wave velocities of peridotites and eclogites despite a difference in density of about 200 kg m^{-3} (see, for example, refs 26, 27). Therefore, we obtain the mantle contribution to PE by satisfying the constraint of local isostasy. The mean elevation of a region in isostatic equilibrium can be divided into crustal (H_c) and mantle (H_m) components²⁰; H_c is directly estimated from the crustal density structure. The difference between H_c and the actual elevation is due to the mantle density structure; this constrains the total mass excess in the mantle lithosphere relative to an equivalent thickness of asthenosphere²⁸.

To calculate the mantle part of the PE we must determine the depth distribution of the mass excess. Many possibilities exist (for example, see ref. 20), but robust constraints are rare. A simple



and flexible assumption is that the mantle density increases proportionately to the n th power of distance above the top of the asthenosphere. In this instance we find that the mantle part of the potential energy is

$$\Delta PE_m = H_m \rho_a g [z_c + L_m / (n + 2)] \quad (6)$$

$$= g z_c \rho_a H_m - g \rho_a^2 H_m^2 / [(n + 2)(\bar{\rho}_m - \rho_a)] \quad (7)$$

We use equation (6) if the thickness of the mantle lithosphere (L_m) is fixed, or equation (7) if the mean density of the mantle lithosphere, $\bar{\rho}_m$, is fixed. For this analysis, we assume a linear ($n = 1$) variation of mantle densities and $\bar{\rho}_m - \rho_a = 50 \text{ kg m}^{-3}$ when H_m is negative. This is a good approximation if mantle densities are determined largely by temperature. If H_m is positive, we place the mass deficit within the mantle above a depth of 100 km and again use $n = 1$ but in equation (6). The resulting ΔPE_m is added to ΔPE_c (equation (4)) to obtain the ΔPE (Figs 1b, 2).

Before interpreting our estimates of ΔPE , it is important to note that the main sources of error in estimating ΔPE (in addition to those related to seismic velocity/density relations²⁸) are the choice of density distribution in the mantle, and the density assignments in the crust. For lithospheric mantle densities that decrease monotonically with depth, distributions of mantle densities that

are consistent both with our estimates of H_m and with plausible densities within the mantle are bracketed by n from about 0 to 2 in equation (7). Because differences in ΔPE due to uncertainties in n are proportional to H_m^2 , uncertainty in ΔPE is greatest (about $7.5 \times 10^{11} \text{ N m}^{-1}$) where H_m is greatest, in the Rockies and Plains (Fig. 1b). Another source of uncertainty is the assignment of a mean density in the mantle lithosphere; variations of the order of 50 kg m^{-3} will create uncertainties in ΔPE increasing with H_m to about $5 \times 10^{11} \text{ N m}^{-1}$ in the Rockies and Great Plains. Both of these sources of uncertainty are apt to be systematic over large regions.

The other principal uncertainty stems from errors in the mantle contribution to elevation, H_m , arising from errors in seismic velocities, density assignments in the crust, or mean elevation. These errors in H_m are less than about 200 m for most points in the Basin and Range, increasing to perhaps 500 m in the Rockies and Plains, where greater flexural thicknesses and poorer constraints on crustal velocities increase potential errors. These uncertainties in H_m produce errors in ΔPE of $< 4 \times 10^{11} \text{ N m}^{-1}$ in the Basin and Range, increasing to perhaps $1.5 \times 10^{12} \text{ N m}^{-1}$ in the Great Plains. Regional ΔPE averages (for example, Fig. 2) will have smaller errors from these factors as they probably are not correlated from measurement to measurement.

An additional uncertainty arises from our choice of the reference state, which we chose to be an asthenospheric column as described above. This produces small positive values of ΔPE at mid-ocean ridges that are consistent with the magnitude of extensional strain rates observed at these very weak parts of the lithosphere. Another reasonable choice would have been to use the global mean PE (as advocated by Coblenz *et al.*²⁹); this would have caused the ΔPE values to be everywhere $1.1 \times 10^{12} \text{ N m}^{-1}$ greater than those shown in Fig. 1c. Thus the zero ΔPE contour would be shifted so as to expand the field of positive ΔPE , but this would not change the relative distribution of PE values, nor would it change qualitatively the calculations

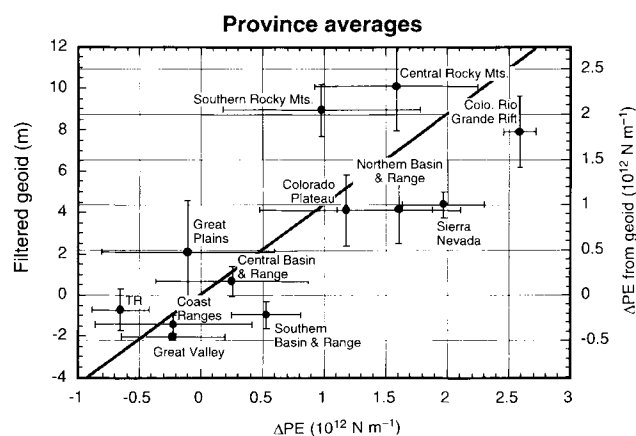


FIG. 2 Comparison of ΔPE from seismic structures and geoid elevations at the same point, averaged over physiographic provinces. Error bars indicate one standard deviation of the population within each physiographic province (Fig. 1a). Geoid values have been converted assuming that 1 m of elevation = $2.281 \times 10^{11} \text{ N m}^{-1}$ of ΔPE (ref. 29). (TR, Transverse Ranges; PR, Peninsular Ranges)

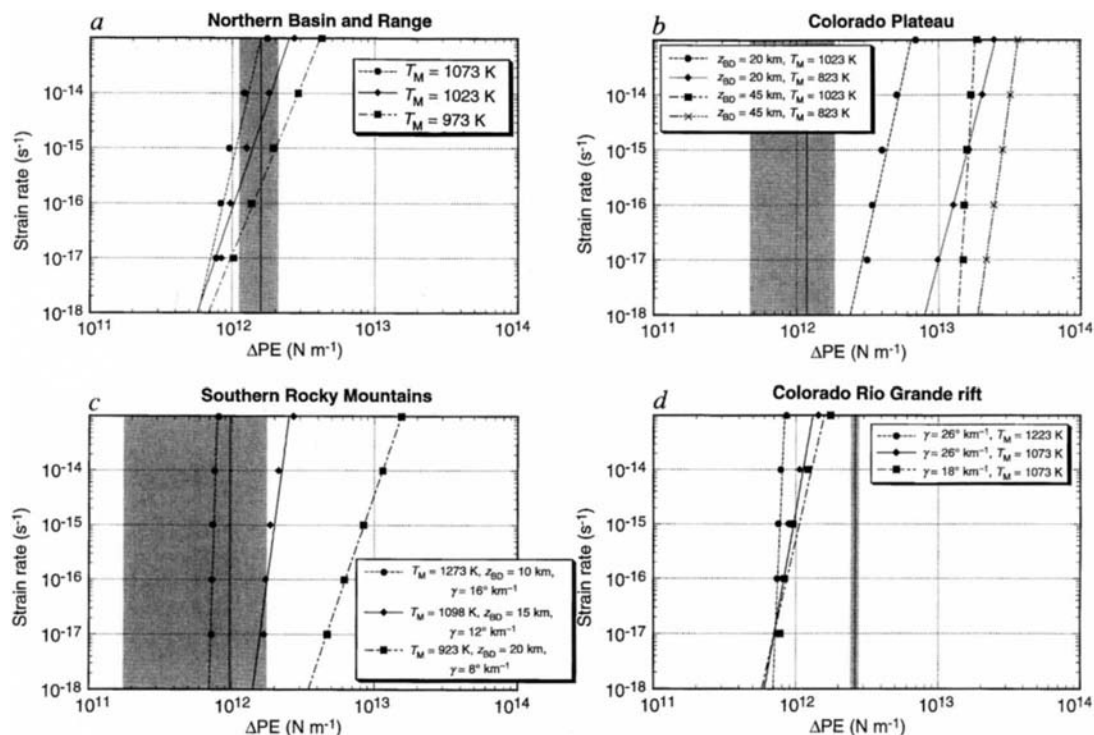


FIG. 3 Vertically averaged horizontal strain rates calculated for tectonic provinces in the western United States. Shaded regions indicate the range of gravitational potential energy available to drive deformation (from Fig. 2). Strain rates were calculated, following Sonder and England³⁸, by assuming a brittle upper-crust rheology governed by Byerlee's law with hydrostatic pore fluid pressure, and a ductile lower crust and mantle governed by creep mechanisms (values of parameters defining these mechanisms are listed in Sonder and England³⁸). Free parameters include crustal thickness (L_c), lithospheric stress state (compressional or extensional), Moho temperature (T_M), depth to brittle–ductile transition (z_{BD}), and mantle thermal gradient (γ). Where the value of parameter is well known, it is given below and is used in all calculations for a tectonic region. Otherwise, values are chosen to reflect the possible range of the parameter, and are indicated on the figure. a, Northern Basin and Range province, extensional stress state. Reduced heat flow of 60 mW m^{-2} (ref. 46) and mantle thermal conductivity

of $3.4 \text{ W m}^{-1} \text{ K}^{-1}$, giving a mantle thermal gradient $\gamma = 18^\circ \text{C km}^{-1}$. $L_c = 30 \text{ km}$; $z_{BD} = 10 \text{ km}$ (refs 47,48). b, Colorado Plateau. Stress state is extensional⁴⁹. $L_c = 45 \text{ km}$ (see Supplementary Information); $z_{BD} = 20 \text{ km}$ is probably a conservative minimum estimate^{49,50}; $\gamma = 9^\circ \text{C km}^{-1}$ (derived from reduced heat flow estimate of Morgan and Gosnold⁴³). c, Southern Rocky Mountain province. Stress state is assumed to be extensional. $L_c = 50 \text{ km}$ (see Supplementary Information). d, Rio Grande Rift. Stress state is extensional^{33,34}. $L_c = 48 \text{ km}$ (see Supplementary Information); $z_{BD} = 10 \text{ km}$ (ref. 51). Parameters based on observations from the Rio Grande rift in New Mexico. As there is more recent vulcanism and extension in New Mexico than in the northward extension of the Rio Grande rift in Colorado, these numbers should be viewed as providing a lower bound on lithospheric strength, and yielding a maximum estimate of strain rates resulting from the available ΔPE . An upper bound is provided by the curves in c.

in Fig. 3. At present, it is difficult from independent evidence to determine the proper reference state, so we must allow that our estimates of ΔPE are probably at the low end of the range of possible values.

Corroboration from the geoid

Theoretically the geoid height should be linearly related to ΔPE (see, for example, ref. 29); however, this relation is limited first by the presence of geoid anomalies associated with features well below the lithosphere (for example, ref. 30), second, by the difficulty of measuring the geoid in continental areas, and third, by certain simplifications inherent to this relation (ref. 29). These difficulties are entirely different from those confronting our estimation of ΔPE from seismic data, so comparison of these two estimates provides some independent validation of our technique.

To reduce the effect of lower-mantle heterogeneities we have removed terms below degree and order 7 (with a cosine taper to degree and order 11; for example, ref. 31) from the GEOID93³² geoid; the result is shown in Fig. 1d. The zero geoid height should be equal to zero ΔPE in the global-mean PE reference frame if the portion of the geoid removed through filtering has no importance for lithospheric potential energy. However, the long-wavelength signal from the difference between oceans and continents suggests that the zero geoid value in Fig. 1d has limited significance. We therefore emphasize the relative differences between different parts of the southwestern United States.

The resulting geoid exhibits patterns quite similar to those in the seismically based image (Fig. 1c). When averaged over physiographic provinces, estimates of ΔPE from the two sets of observations are quite similar (Fig. 2). The differences that exist are probably caused by errors in one or both of the techniques for obtaining ΔPE as discussed above. Although further examination of these differences will be fertile territory for better understanding the structure of the lithosphere, the overall validation of the seismically based ΔPE measurements by the geoid observations justifies a discussion of the tectonic implications of these anomalies.

Potential energy and active deformation

Most broadly, the zero value of seismically based ΔPE in Figs 1c and 2 divides extending from non-extending lithosphere. The most positive ΔPE s are not in the regions of highest topography, but instead are within the somewhat lower, but actively extending, Great Basin and the Rio Grande rift. Values near zero in the southern US Basin and Range compare well with the low extensional strain rates interpreted for that region. Negative ΔPE s in the Great Plains are consistent with the compressive stresses reported there^{33,34}. Contractional deformation in coastal California (Coast Ranges, Transverse Ranges)^{35–37} is consistent with the negative ΔPE s found there.

We can attempt to test quantitatively the general agreement of ΔPE and present-day strain rates by applying PE-derived stresses

to the lithosphere and calculating the resulting strain rates. The expected vertically averaged horizontal strain rate for a given ΔPE is highly dependent on assumptions about the strength of the lithosphere, which in turn depend on rheology, thermal state, pore fluid pressure and the stress regime. But it is possible to make reasonable assumptions about such parameters based on previous geological and geophysical investigations in the Cordillera. Following Sonder and England³⁸ and using values of parameters appropriate to the Great Basin, the ΔPE of the Great Basin should result in extensional strain rates between 10^{-14} and 10^{-16} s^{-1} (Fig. 3a). Similar rates are predicted by the models of Kusznir and Park³⁹ and Lynch and Morgan⁴⁰. These predicted rates are compatible with Eddington *et al.*'s⁴¹ geodetically and seismically observed deformation rates of approximately 10^{-15} to 10^{-16} s^{-1} across the Great Basin.

In the Colorado Plateau, the calculated potential energy is not capable of driving geologically significant deformation (Fig. 3b), consistent with the negligible strain rates observed today⁴². In the Rocky Mountains, observations relevant to constraining lithospheric rheology are highly variable. Reduced heat flux estimates, which can constrain mantle temperature gradients, vary by a factor of two^{43,44}. Moreover, this variability probably reflects differences in crustal heat sources⁴⁴, which makes estimates of the thermal gradients in the mantle all the more uncertain. The maximum depth of crustal seismicity, which can be used to constrain the depth to the brittle–ductile transition, is poorly constrained in the Rocky Mountains owing to both sparse seismicity and infrequent observations. Nevertheless, for all reasonable choices of parameters the Rocky Mountain lithosphere is too strong to deform unless the mantle has negligible strength and lower-crustal temperatures approach $1,000^\circ\text{C}$ (Fig. 3c). The only significant strain rates in the Rocky Mountain region are localized along the Rio Grande rift and its northward extension in Colorado. This axis contains the only notable late Cenozoic volcanic eruptions, and thus probably is the only part of the Rocky Mountains with mantle temperatures high enough to permit buoyancy-driven deformation. Indeed, our estimates of ΔPE are high enough to produce relatively high strain rates (Fig. 3d); the lower observed rates probably reflect the influence of boundary forces from adjacent regions on this long, narrow province.

These results demonstrate that the PE distribution in the Cordillera is sufficient to produce the observed present-day

deformation, or lack thereof. In the Great Basin, the lithosphere is sufficiently weak that stresses arising from PE differences are capable of producing geologically significant strain rates. We infer a similar relationship for the actively extending Rio Grande rift. In contrast, the lithosphere is sufficiently strong in the Colorado Plateau and Rocky Mountains to be able to resist the stresses arising from isostatically supported topography and horizontal density contrasts. Our analysis has quantified the stresses arising from buoyancy in the western United States, and by so doing, has shown that it is not necessary to invoke external forces arising from far-field plate kinematics⁹ or tractions at the base of the lithosphere^{4,6,45} to account for active deformation of the interior of the Cordillera.

Origins of lithospheric ΔPE

One advantage of the seismically based approach to calculating ΔPE is that the relative contributions of the crust and mantle may be distinguished. For the southwestern United States we can compare Fig. 1b and c to see that potential energy contrasts have arisen from different sources in different places. In the Basin and Range, the high values of H_m suggest that a very buoyant mantle is responsible for the current extensional deformation of this region. In contrast, the mantle is much less buoyant in the Rocky Mountains; much of the potential energy in this region has been stored in the crust, probably during the Early Tertiary Laramide orogeny. Within the southern Rockies, only the Rio Grande rift exhibits significant late Cenozoic deformation; this deformation appears to be directly related to the locally higher values of H_m under the Rio Grande rift through either or both an increase in ΔPE from the mantle and a change in rheology from mantle heating.

The ability to isolate and contrast contributions of the mantle and crust to potential-energy-based deformation of the lithosphere is a strength of our approach, and potentially allows extrapolation of this technique into the past. Geological constraints on changes in crustal structure and deformational style through time can be used to infer probable contributions of the crust to ΔPE . Deformational style and constraints on mean elevation can limit the possible distributions of mass within the upper mantle. This quantitative analysis can thus test presently speculative models explaining the tectonic history of the Cordillera and other regions. □

Received 15 February; accepted 2 April 1996.

- Christiansen, R. L. & Yeats, R. S. *The Geology of North America (DNAG)* Vol. **G-3**, 261–406 (Geol. Soc. Am., Boulder, Colorado, 1992).
- Braile, L. W., Hinze, W. J., von Frese, R. R. B. & Keller, G. R. *Geol. Soc. Am. Mem.* **172**, 655–680 (1989).
- Eaton, G. P. A. *Rev. Earth planet. Sci.* **10**, 409–440 (1982).
- Gough, D. I. *Nature* **311**, 428–432 (1984).
- Bird, P. *J. geophys. Res.* **94**, 3967–3990 (1989).
- Bird, P. *Science* **239**, 1501–1507 (1988).
- Wise, D. U. *Geol. Soc. Am. Bull.* **74**, 357–362 (1963).
- Atwater, T. M. *Geol. Soc. Am. Bull.* **81**, 3518–3536 (1970).
- Ingersoll, R. V. *Geology* **10**, 621–624 (1982).
- Molnar, P. & Chen, W.-P. *J. geophys. Res.* **88**, 1180–1196 (1983).
- Sonder, L. J., England, P. C., Wernicke, B. P. & Christiansen, R. L. in *Continental Extensional Tectonics* (eds Coward, M. P., Dewey, J. F. & Hancock, P. L.) 187–201 (Spec. Publ. 28, Geol. Soc., Oxford, 1987).
- Coney, P. J. & Harms, T. A. *Geology* **12**, 550–554 (1984).
- Molnar, P. & Lyon-Caen, H. *Geol. Soc. Am. Spec. Pap.* **218**, 179–207 (1988).
- Dewey, J. F. *Tectonics* **7**, 1123–1139 (1988).
- Platt, J. P. & Vissers, R. L. M. *Geology* **17**, 540–543 (1989).
- Zhou, S. & Sandiford, M. *J. geophys. Res.* **97**, 14207–14221 (1992).
- Artyushkov, E. V. *J. geophys. Res.* **78**, 7675–7708 (1973).
- Turcotte, D. L. *J. geol. Soc. Lond.* **140**, 701–724 (1983).
- Houseman, G. & England, P. C. *J. geophys. Res.* **91**, 719–729 (1986).
- Lachenbruch, A. H. & Morgan, P. *Tectonophysics* **174**, 39–62 (1990).
- Christensen, N. I. & Mooney, W. D. *J. geophys. Res.* **100**, 9761–9788 (1995).
- Ludwig, W. J., Nafe, J. E. & Drake, C. L. in *New Concepts of Sea Floor Evolution, Pt 1, Regional Observations* (ed. Maxwell, A. E.) 53–84 (Wiley Interscience, New York, 1970).
- Brothie, J. F. & Silvester, R. *J. geophys. Res.* **74**, 5240–5252 (1969).
- Bechtel, T. D., Forsyth, D. W., Sharpston, V. L. & Grieve, R. A. F. *Nature* **343**, 636–638 (1990).
- Birch, F. *J. geophys. Res.* **66**, 2199–2224 (1961).
- Kern, H. M. *Phys. Earth planet. Inter.* **79**, 113–136 (1993).
- Anderson, D. L. *Theory of the Earth* (Blackwell Scientific, Boston, 1989).
- Jones, C. H. *et al. Tectonophysics* **213**, 57–96 (1992).
- Coblentz, D. D., Richardson, R. M. & Sandiford, M. *Tectonics* **13**, 929–945 (1994).
- Hager, B. H., Clayton, R. W., Richards, M. A., Comer, R. P. & Dziewonski, A. M. *Nature* **313**, 541–545 (1985).

- Sheehan, A. F. & Solomon, S. C. *J. geophys. Res.* **96**, 19981–20009 (1991).
- Milbert, D. G. in *Gravity CD-ROM and User's Manual* (eds Hittelman, A. M., Dater, D. T., Buhmann, R. W. & Racey, S. D.) (National Oceanic and Atmospheric Administration, National Geophysical Data Center, Boulder, CO, 1994).
- Zoback, M. D. & Zoback, M. L. *Decade Map Vol. 1*, 339–366 (Geol. Soc. Am., Boulder, Colorado, 1991).
- Zoback, M. L. *et al. Decade of N. Am. Geol. Continent-Scale Map 5*, (Geol. Soc. Am., Boulder, Colorado, 1991).
- Unruh, J. R., Loewen, B. A. & Moores, E. M. *Geol. Soc. Am. Bull.* **107**, 38–53 (1995).
- Wakabayashi, J. & Smith, D. L. *Bull. seism. Soc. Am.* **84**, 1960–1970 (1994).
- Humphreys, E. D. & Weldon, R. J. II. *J. geophys. Res.* **99**, 19975–20010 (1994).
- Sonder, L. J. & England, P. C. *Earth planet. Sci. Lett.* **77**, 81–90 (1986).
- Kusznir, N. J. & Park, R. G. in *Continental Extensional Tectonics* (eds Coward, M. P., Dewey, J. F. & Hancock, P. L.) 35–52 (Spec. Publ. 28, Geol. Soc., Oxford, 1987).
- Lynch, H. D. & Morgan, P. in *Continental Extensional Tectonics* (eds Coward, M. P., Dewey, J. F. & Hancock, P. L.) 53–65 (Spec. Publ. 28, Geol. Soc., Oxford, 1987).
- Eddington, P. K., Smith, R. B. & Renggli, C. in *Continental Extensional Tectonics* (eds Coward, M. P., Dewey, J. F. & Hancock, P. L.) 371–392 (Spec. Publ. 28, Geol. Soc., Oxford, 1987).
- Menges, C. M. & Pearthree, P. A. *Ariz. geol. Soc. Digest* **17**, 649–680 (1989).
- Morgan, P. & Gosnold, W. D. *Geol. Soc. Am. Mem.* **172**, 493–522 (1989).
- Decker, E. R., Heasler, H. P., Buelow, K. L., Baker, K. H. & Hallin, J. S. *Geol. Soc. Am. Bull.* **100**, 1851–1885 (1988).
- Bird, P. *Tectonics* **3**, 741–758 (1984).
- Lachenbruch, A. H. & Sass, J. H. *Geol. Soc. Am. Mem.* **152**, 209–250 (1978).
- Meissner, R. & Strehlau, J. *Tectonics* **1**, 73–90 (1982).
- Sibson, R. H. *Bull. seism. Soc. Am.* **72**, 151–163 (1982).
- Wong, I. G. & Humphrey, J. R. *Geol. Soc. Am. Bull.* **101**, 1127–1146 (1989).
- Wong, I. G. & Chapman, D. S. *Bull. seism. Soc. Am.* **80**, 589–599 (1990).
- Jaksha, L. H., Locke, J. & Gebhart, H. J. *Geol. Soc. Am. Bull.* **92**, 31–36 (1981).
- Isaaks, E. H. & Srivastava, R. M. *An Introduction to Applied Geostatistics* (Oxford Univ. Press, 1989).

SUPPLEMENTARY INFORMATION. Available on Nature's World-Wide Web site <http://www.nature.com>. Paper copies are available from Mary Sheehan at the London editorial office of Nature.

ACKNOWLEDGEMENTS. This Article was improved by thoughtful reviews from A. Lowry and P. C. England. C.H.J. was supported by the Geophysics Program of the NSF. L.J.S. received support from the Dartmouth College Faculty Research and Development Fund. J.R.U. acknowledges support for professional development from William Lettis & Associates, Inc.

Structure and mutational analysis of Rab GDP-dissociation inhibitor

Isabelle Schalk^{*†‡}, Ke Zeng^{*†}, Shih-Kwang Wu[†], Enrico A. Stura^{*}, Jeanne Matteson[†], Mingdong Huang^{*}, Anurag Tandon[†], Ian A. Wilson^{*} & William E. Balch[†]

Departments of ^{*} Molecular and [†] Cell Biology, The Scripps Research Institute, 10666 N. Torrey Pines Road, La Jolla, California 92037, USA

The crystal structure of the bovine α -isoform of Rab GDP-dissociation inhibitor (GDI), which functions in vesicle-membrane transport to recycle and regulate Rab GTPases, has been determined to a resolution of 1.81 Å. GDI is constructed of two main structural units, a large complex multisheet domain I and a smaller α -helical domain II. The structural organization of domain I is surprisingly closely related to FAD-containing monooxygenases and oxidases. Sequence-conserved regions common to GDI and the *choroideraemia* gene product, which delivers Rab to catalytic subunits of Rab geranylgeranyltransferase II, are clustered on one face of the molecule. The two most sequence-conserved regions, which form a compact structure at the apex of GDI, are shown by site-directed mutagenesis to play a critical role in the binding of Rab proteins.

THE specific membrane association and recycling of the Rab family of GTPases¹, which are critical in vesicle-membrane trafficking², are regulated by Rab GDP-dissociation inhibitor (GDI). α -GDI, the prototype member of a ubiquitous family of highly related proteins, forms a complex with the GDP-bound, prenylated form of Rab3A (ref. 3) and can interact with and retrieve a broad range of Rab proteins from the membrane bilayer⁴, being a key regulator of their function. The heterodimeric complex serves as a cytosolic reservoir for the recycling of inactive Rab for use in multiple rounds of vesicle budding and fusion¹. GDI family members are significantly related through sequence-conserved regions (SCRs) to the mammalian *choroideraemia* gene product CHM, which contributes to late-onset retinal degeneration when defective^{5–10}. CHM and a related isoform, CHML, are identical to the component-A subunit of Rab geranylgeranyl transferase II (refs 5, 7, 8, 11), also referred to as Rab-escort protein (REP), which delivers newly synthesized Rab to the catalytic subunits for prenylation^{5,7,8,11–13}. After prenylation, REP function is indistinguishable from that of GDI¹⁴.

We report here the three-dimensional structure of the brain α -GDI isoform. An unexpected finding was that the structure of GDI is remarkably similar to flavoproteins that function as monooxygenases¹⁵ or oxidases^{16,17}. Two of the SCRs form a compact structure referred to as the GDI *choroideraemia* consensus domain (GCD). We demonstrate by site-directed mutagenesis that conserved residues in the GCD fold to form a Rab-binding region at the apex of GDI.

Structure description

The X-ray structure was determined at 1.81 Å resolution by multiple isomorphous replacement (MIR) (Table 1 and Fig. 1). GDI appears to have two major domains, a large cylindrical module consisting of ~340 residues and a smaller, ~100-residue α -helical domain (Fig. 1b). The large domain contains two major β -sheets, one parallel (α 1– α 5) and the other mostly antiparallel (β 1– β 7), which dominate the core of the molecule (Fig. 1b, c). The two sheets are nearly perpendicular to each other and are connected by two long β -ribbon-like strands (β 3 and β 7). Two additional antiparallel β -sheets, c and e, are flanked on the

external side of sheet a and the internal side of sheet b, respectively. The overall structure is composed of five main layers in packing order: sheet c, sheet a, three parallel helices (I, A, N), sheets e and f, and finally the largest sheet b, traversed externally by a short helix (M). An additional helix C runs perpendicular to the other three parallel helices I, A and N, and completes a largely hydrophobic packing of the four helices. The packing of sheet a against the I, A, N helices is also predominantly hydrophobic.

The smaller helical domain II (residues 119–219) consists of five helices folded as an independent unit between the most external strand b2 of sheet b and the central strand e3 of sheet e (Fig. 1b, c). Three antiparallel helices E, F and G form a tripod-like structure with helix H as the base. A hydrophobic core is formed between these helices by nine aromatic residues (Phe 140, 143, 146, 158, 184; Tyr 172, 192, 197). These two major domains (I and II) form a 'V' shape with a large groove bordered by helix H of the small domain and the C-terminal helix N of the large domain with β -sheet b lying at its base (Fig. 1b).

GDI compared with similar structures

The program DALI (version 2.0)¹⁸ was used to search for structurally similar proteins among 642 structures in the Protein Data Bank. The top three of 46 aligned structures showed significant similarity to GDI in their three-dimensional folding as well as in their topological connections. *p*-Hydroxybenzoate hydroxylase (PBHase)^{19,20}, cholesterol oxidase (COX)¹⁷ and glucose oxidase (GOX)¹⁶ had alignment z-scores of 15.6, 12.2 and 11.2, respectively, with almost identical topologies for domain I (shown for PBHase in Fig. 2a). The alignment is particularly prominent in β -sheets a and c, and in helices I, A, N and C. GDI and PBHase differ in three regions: strands f1, f2 and the B helix in domain I (Fig. 2b, dark green) and all of domain II (Fig. 2b; blue) in GDI are absent in PBHase, whereas PBHase contains an extended carboxy terminus (Fig. 2c, light green).

PBHase, COX and GOX are a structural family of FAD-dependent flavoproteins. Bound FAD occupies a long groove formed between the five-stranded parallel sheet a and the seven-stranded antiparallel sheet b^{16,17,20} (Fig. 2c). The ADP moiety of FAD is bound next to a GxGxxG motif (x represents non-glycine residues) located between the first β -strand a1 and α -helix A of the amino-terminal $\beta\alpha\beta$ unit (Fig. 2c). This motif^{15,21} has been shown to be a consensus for phosphate binding in many nucleotide-

[†] Present address: Laboratoire de Chimie Microbienne, Institut de Chimie, BP 296 R/S, 1 rue Blaise Pascal, 67008 Strasbourg Cedex, France.

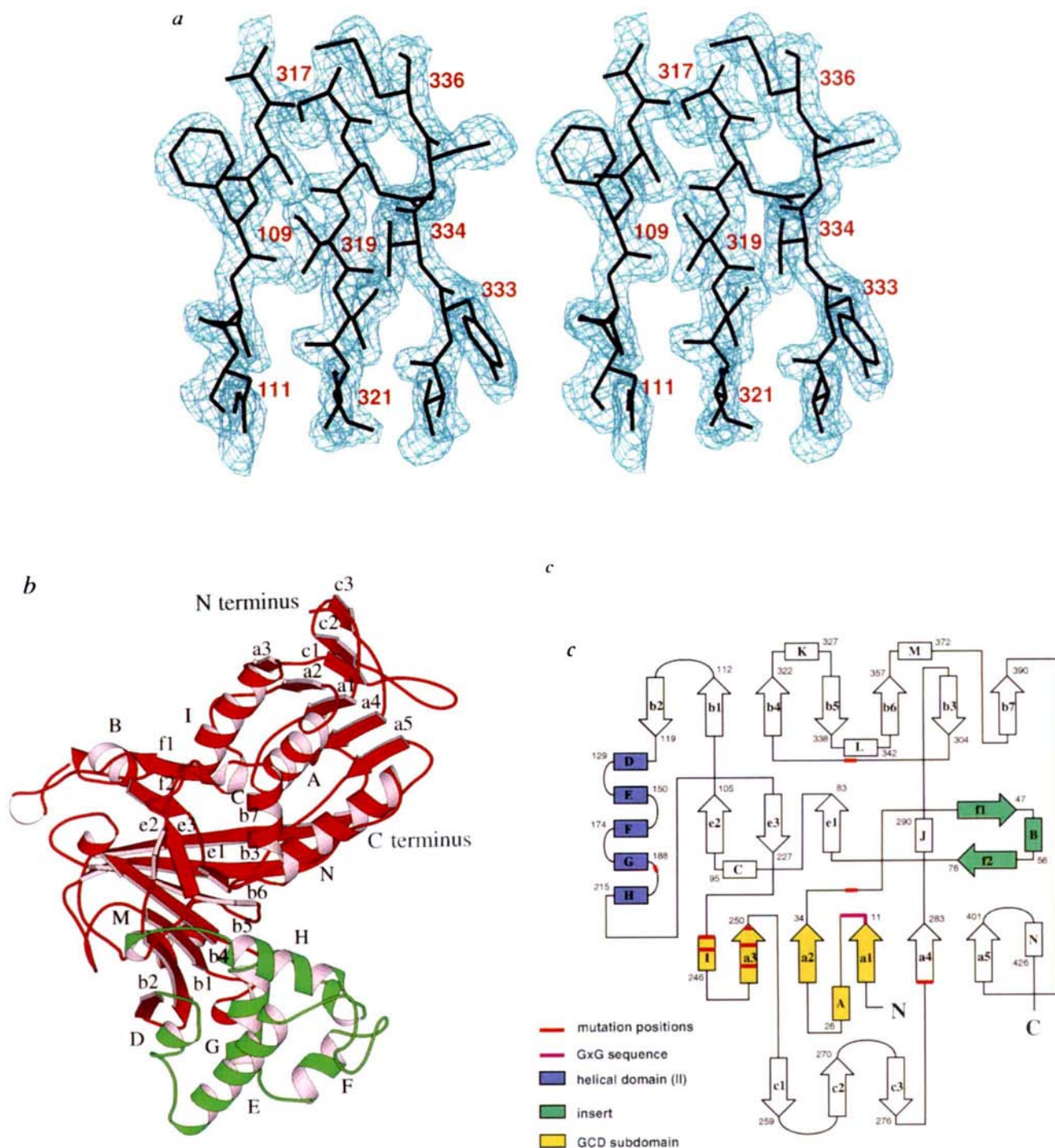


FIG. 1 *a*, Stereo view of $2F_o - F_c$ electron-density map at 1.81 Å resolution contoured at 1σ for β -sheet *b* (strands *b1*, *b4* and *b5* only). The 'omit' map was calculated by adding a random error (± 0.25 Å) to the refined model coordinates while the shown part was omitted from the map calculation. *b*, Stereo pair of a $C\alpha$ ribbon representation of GDI structure drawn with MOLSCRIPT⁴². The two domains I and II, are shown in red and green, respectively, with the overall dimensions of $\sim 60 \times 40 \times 40$ Å. The secondary structure assignment was determined using DSSP⁴³. Helices are labelled alphabetically in capital letters according to their appearance in the sequence; β -sheets are labelled in lower case, with numbers designating their order in the sequence. The alphabetical labels for the β -sheets are consistent with those of PBHase¹⁵ except for additional strands *f1* and *f2* in GDI. The sheet designated as 'd' in PBHase was not assigned in GDI owing to the lack of convincing β -strand interactions, although there are equivalent

extended strands in a similar location. The N and C termini are indicated. The small 3_{10} helices J, K and L are not labelled. *c*, Topological connections for the GDI structure. Arrows and rectangles represent β -strands and α -helices, respectively. Strands that form β -sheets are grouped together. Some residue numbers are indicated for the locations of the secondary structural elements. The green and blue 'inserts' are not found in PBHase (Fig. 2). The sequence GxG in the first $\beta\alpha\beta$ unit is coloured magenta and aligns with the corresponding sequence in PBHase. Site-directed mutagenesis was carried out at sequence sites Y39, D195, E233, R240, T248, T249, M250, K278 and K309, which are indicated by red bars. The GDI CHM consensus domain (GCD) comprising SCRs 1A, 1B and 3B is shown in yellow and includes three strands and two helices, including the first $\beta\alpha\beta$ unit containing the GxG sequence.

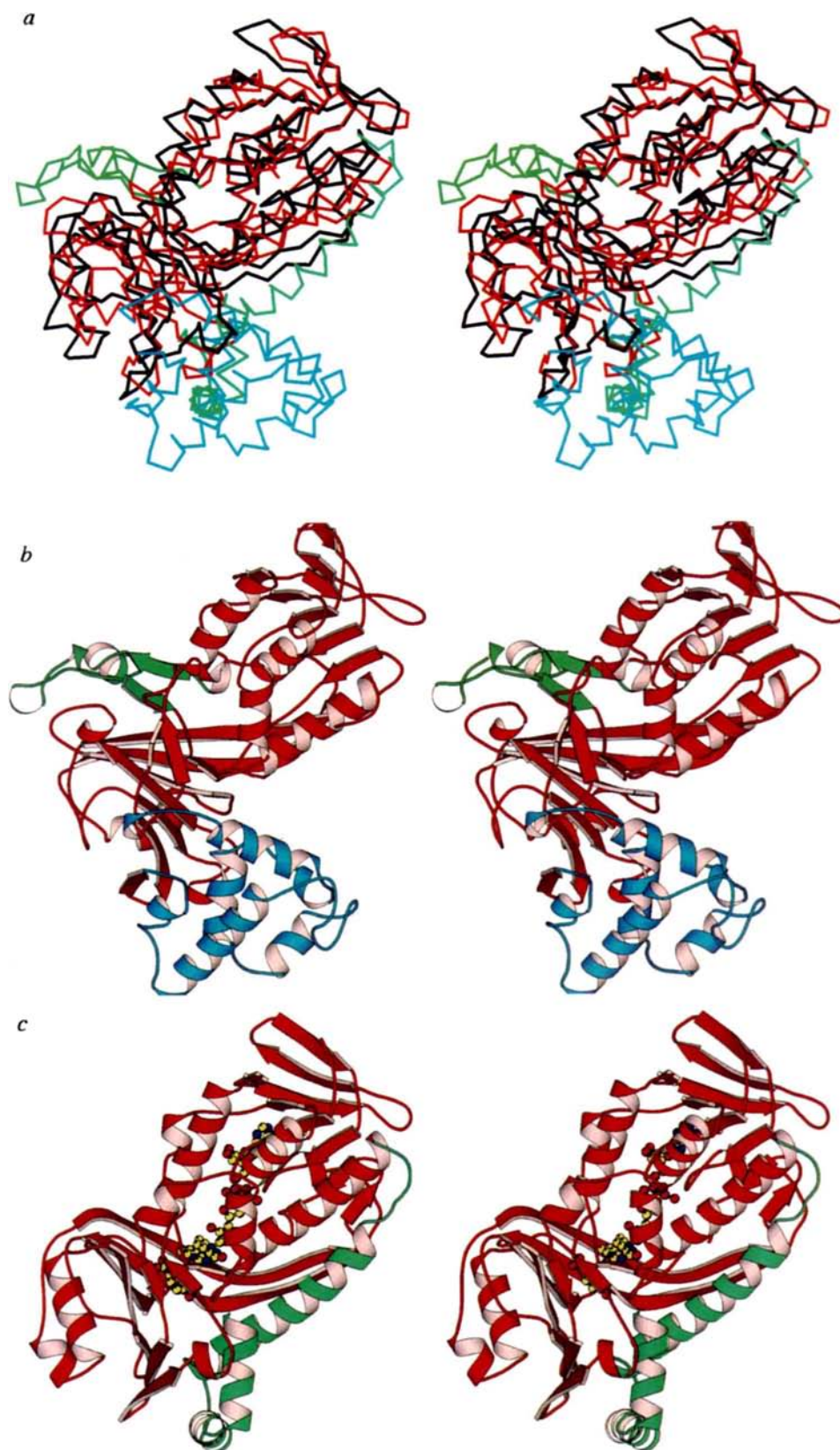


FIG. 2 *a*, Superposition of the C α atoms of GDI (red) and PBHase (black) based on the distance matrix alignment method¹⁸. A total of 253 structurally equivalent C α atoms were aligned with an overall r.m.s.d. of 3.9 Å. All of the aligned segments have the same chain topological connections. The two major inserts in GDI compared to PBHase are shown in dark green (strands f1, f2, and helix B) and blue (domain II), and the C-terminal α -helical extension of PBHase is shown in light green. *b*, *c*, Stereoview of the three-dimensional structures of GDI (*b*) and PBHase (*c*) drawn in MOLSCRIPT⁴². Local structure comparison between GDI and PBHase yields a 0.99-Å r.m.s. deviation for the first $\beta\alpha\beta$ unit, consisting of strand a1-helix A-strand a2 (36 C α atoms), 0.72 Å for sheet a (26 C α atoms), 0.92 Å for sheet c (21 C α atoms) and 2.1 Å for sheet b (43 C α atoms). In the comparison of GDI (*b*) and PBHase (*c*), similar structural regions are drawn in the same colour (red), and differences in inserts are indicated by different colours. The helical domain II of GDI can easily be accommodated as an insert in PBHase between strands b2 and e3 (*b*, blue), whereas the last three helices at the C terminus of PBHase are absent in GDI (*c*, light green). In addition, the two-stranded sheet f (f1 and f2) and helix B in GDI form an extra insert between strands a2 and e1 (*b*, dark green) that is absent in PBHase. The location of the FAD cofactor (ball-and-stick representation) is shown in *c*.

binding proteins¹⁶. All members of the GDI family contain a GxG remnant of this motif (residues 11–13) which structurally aligns (Fig. 1c) with the corresponding GxGxxG sequence of PBHase (Fig. 2a). In experiments designed to explore the unexpected structural homologies between GDI and the flavoproteins, we have been unable to detect flavin binding either to recombinant GDI prepared from *Escherichia coli* or *Sporidia frugiperdia* (SF9) cells, or to α -GDI freshly prepared from bovine brain³. Although soaking of the GDI crystals with FAD led to a brilliant yellow coloration, a hallmark of flavoproteins²², association appeared to be nonspecific in that the structure of soaked crystals showed no

additional density in the region homologous to PBHase (Fig. 2c)²⁰. Thus the functional significance of the strong structural similarities, although provocative, is unclear.

GDI/CHM conserved domain

GDI is highly conserved throughout evolution and this conservation extends to members of the *choroideraemia* gene family (Fig. 3a). Sequence alignment of CHM and GDI shows five SCRs located in the N-terminal and central portions of the molecule. Residues comprising motifs found in SCRs 1A, 1B and 3B are particularly diagnostic of GDI and CHM family members owing to

TABLE 1 Statistics of data collection and structure refinement

Data set	Resolution (Å)	Completeness (%)	R_{sym} (%) [*]	R_{iso} (%) [†]	Sites	R_{cullis} (%) [§]	Phasing power (iso/ano) [‡]	Figure of merit (iso/ano)
Native	1.81	95 (83)	10.0 (32)	—	—	—	—	—
EtHgCl	2.7	75	3.0	12.7	4	53	2.52/1.95	0.46/0.31
EMTS	2.91	60	6.8	15.8	6	55	2.30	0.40
UO ₂ Ac1	2.7	92	6.6	9.4	5	62	1.69/1.37	0.36/0.31
UO ₂ Ac2	2.6	69	8.1	14.6	3	58	1.38	0.34
								Overall 0.78
Refinement	Resolution range	No. reflections		No. protein atoms	No. water molecules		R -value (%)	R_{free} (%)
	6.0–1.81 Å	33,282 (40,982) [#]		3,409	410		19.8 (21.1) [#]	27.5 (30.2) [#]
R.m.s. deviation	Bond lengths	Bond angles		Dihedrals		Impropers		Coordinate error (from Luzzati plot)
	0.010 Å	1.66°		25.2°		1.42°		0.20–0.25 Å

Crystallization and data collection. Native GDI crystals were obtained from ammonium sulphate (1.73 M, pH 6.25) in the presence of L- α -phosphatidylcholine dimyristol using reverse screening and macroseeding methods as described^{30,31}. The crystals belong to monoclinic space group $P2_1$ with cell dimensions $a = 91.9$ Å, $b = 43.5$ Å, $c = 63.2$ Å and $\beta = 104.5^\circ$. The crystals have a Matthews coefficient³² of 2.18 Å³ per dalton which corresponds to one molecule of M_r 55K in the asymmetric unit. Two usable mercury derivatives were prepared by soaking native crystals for 24 h in mother liquors containing 0.02 mM ethylmercury chloride (EtHgCl) and 0.05 mM ethylmercurithiosalicylic acid (EMTS), respectively. The two derivatives share four common and nearly overlapping sites, with the EMTS derivative possessing an additional two sites. A three-site uranyl acetate derivative (UO₂Ac2) was obtained by soaking native crystals in mother liquor containing 1.4 mM uranyl acetate for 48 h. By increasing the uranyl concentration to 9 mM and reducing the soaking time to 24 h, two additional sites were obtained in a second uranyl derivative (UO₂Ac1). The native X-ray data set to 1.81 Å resolution was collected on a MAR image plate whereas all the derivative data sets were collected on a Siemens X1000 area detector using double-mirror-focused CuK α radiation from a Siemens rotating anode³³. Anomalous data for all the derivatives were also collected, but only UO₂Ac1 and EtHgCl derivatives gave significant anomalous signals. Data were processed using DENZO (Z. Otwinowsky) and XENGEN (v2.1)³⁴ program packages. **Patterson solution and phase calculation.** Isomorphous difference Patterson maps at 4.0 Å resolution provided clear solutions for two sites for both EtHgCl and UO₂Ac1 derivatives. The other sites were readily revealed from difference Fourier maps using phases from the known sites. Cross-difference Fourier maps were also used in searching for minor sites and to test the consistency of the handedness of the solutions from the various crystals. Anomalous difference Patterson maps for EtHgCl and UO₂Ac1 derivatives yielded sites consistent with those from the isomorphous difference Patterson functions. Heavy-atom parameters were refined using XtalView³⁵ and PHASIT in PHASES³⁶ package. Initial MIR phases at 3.0 Å resolution using both isomorphous and anomalous data were obtained with solvent-flattening³⁷ procedures as implemented in PHASES. A molecular mask was built assuming a solvent content of $\sim 40\%$ in the unit cell. The resultant electron density map, with a mean figure of merit of 0.78, clearly showed the molecular boundary and allowed tracing of about 400 α -carbons of the 447 residues. The only Trp residue (Trp 72) was identified from its side-chain density as well as from its relative location in the model trace. Phase extension to 2.8 Å using combined MIR and C α partial model phases still resulted in broken density for many loop regions on the protein surface, and no obvious density for the ~ 40 missing residues. Phases were recalculated with solvent flattening using molecular masks constructed from the traced model plus a 10-Å cushion enlarged from its surface in an attempt to recover the missing residues. This latter model was only used for mask construction and was not used in the MIR phasing. The topological connections of the C α tracing was confirmed through this process and the electron density was improved in previously ambiguous loop regions. The current model contains 430 residues of the total 447, with 17 not observed at the carboxyl end. The six Hg atoms of derivatives were all found to be close to 5 of the total 10 free cysteine residues. Two of the six Hg atoms appeared to be very close to a single cysteine residue. **Refinement.** The model was refined using the slow-cooling protocol in XPLOR (version 3.1)³⁸ and rebuilt against 'shake-omit' $2F_o - F_c$ maps generated with a random error (± 0.25 Å) added to the refinement model³⁵ for segments of $\sim 8\%$ of the molecule omitted. A number of regions were corrected and improved on the basis of this procedure. Solvent water molecules were identified using WAT_PCKPIK routine in O (ref. 39) and were checked with cross-validation R_{free} ⁴⁰ during the process. The final crystallographic R -value and R_{free} (based on a data set of 10% randomly omitted observations sequestered before any refinement) are 0.198 and 0.275, respectively, for 33,282 independent reflections from 1.81 to 6.0 Å resolution ($F > 2\sigma(F)$). The overall B values (refined isotropically) for the 3,409 non-hydrogen and water oxygen atoms are 28.2 and 30.8 Å², respectively. The overall real space fit correlation against $2F_o - F_c$ maps calculated with the program O is 0.91. A portion of β -sheet b fitted into the $2F_o - F_c$ 'shake-omit' map is shown in Fig. 1. A Ramachandran plot (not shown) generated from PROCHECK⁴¹ showed that over 86% of main-chain torsion angles are in most favoured regions, with only Lys 264 and Glu 395 in 'disallowed' regions. These two residues are located in loop regions on the surface and have poor densities, low real-space fit correlation coefficients and high thermal parameters, indicating positional disorder in these regions. The average crystallographic thermal parameters for domains I and II are 25.7 and 28.7 Å², respectively.

^{*} $R_{\text{sym}} = \sum_n \sum_i |I_i(h) - \langle I(h) \rangle| / \sum_n \sum_i I_i(h)$, where h is the index for unique reflections, $I_i(h)$ is the i th measurement of reflection h and $\langle I(h) \rangle$ the mean intensity of all measurements of $I(h)$.

[†] $R_{\text{iso}} = \sum_n |F_{\text{nat}} - F_{\text{deriv}}| / \sum_n |F_{\text{nat}}|$, where F_{nat} and F_{deriv} are the structure factor amplitudes for native and derivative, respectively.

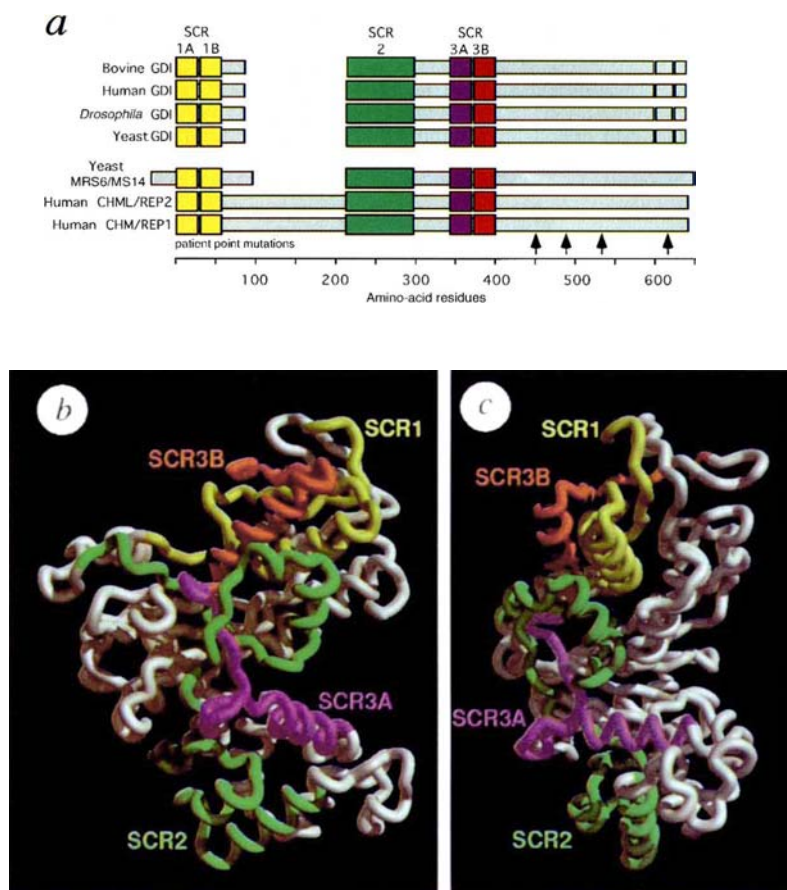
[‡] Phasing power = F_H/E , where F_H and E are the r.m.s. of heavy-atom structure factor and lack-of-closure error, respectively.

[§] $R_{\text{cullis}} = \sum_n |F_{\text{deriv}} - F_{\text{nat}}| - |F_{\text{Hcalc}}| / \sum_n |F_{\text{deriv}} - F_{\text{nat}}|$, where the summations are over centric reflections, F_{Hcalc} is the calculated structure factor for heavy atoms.

^{||} Outer shell (1.81–1.98 Å) for $I \geq 0$.

[#] $F > 2\sigma(F)$ (all data).

FIG. 3 Distribution of SCRs in GDI and CHM family members. **a**, Sequence-conserved regions (SCRs)¹⁰. The N-terminal half of GDI shares conserved regions (SCRs 1–3) with members of the CHM family. Human CHM/CHML contain a large insertion (138 amino acids) relative to GDIs or yeast MRS6/MS14 that separates SCR1B and SCR2. Vertical arrows indicate the locations of some of the nonsense and frameshift mutations observed in choroideremia patients⁴⁴. Boxes correspond to SCRs common to GDI and CHM, and are colour-coded to match equivalent residues in the GDI structure (**b**, **c**): SCR1A/1B (yellow), SCR2 (green), SCR3A (purple) and SCR3B (red). The consensus sequences found in evolutionarily divergent species are as follows: DVxxxGTGxxExxL (SCR1A), GxxVLHxDxxxYYG (SCR1B) and GExxQGFxRxxAxxG (SCR3B). These can be used to identify uniquely members of GDI and CHM families¹⁰. Accession numbers for the sequences used are: bovine α -GDI (GenBank D90103), human α -GDI (D13988), *Drosophila* GDI (GenBank L03209), yeast MRS6/MS14 (GenBank X70339), human CHML/REP2 (GenBank X64728) and human CHM/REP1 (GenBank X78121). **b**, **c**, Backbone tube representation of GDI showing the location of SCRs drawn using *MidasPlus* version 2.0 (ref. 45). In **b**, the molecular orientation of the GDI structure is similar to that shown in Fig. 1; colour-coding is the same as in **a**. SCR1 (yellow; residues 4 to 47), and SCR3B (red; residues 232 to 259) from the compact GCD subdomain found at the apex of domain I. SCR3A (purple; residues 203 to 230) and SCR2 (green; residues 69 to 160) extend along the face connecting domains I and II. In **c**, GDI is rotated by 90° around the vertical axis compared to **b**. Note that the SCRs are largely confined to one side of the molecule, suggesting that this face is involved in Rab association and/or interaction with surface membrane receptors.



their composition of invariant di- and tripeptides in evolutionarily divergent species¹⁰.

Strikingly, the residues found in SCRs 1 and 3B form a compact structure at the apex of the molecule. This GDI–CHM consensus domain (GCD) encompasses strand a1, helix A, and strand a2 of β -strand f1 in the extreme N-terminal region at the sequence (Fig. 3b, yellow). Despite being separated by nearly 140 residues, SCR3B winds back to contact SCR1 as the outermost strand a3 of sheet a (Fig. 3b, red). The less conserved regions SCR2 (green) and SCR3A (purple) also show significant interaction and extend from the lower half of domain I into domain II (Fig. 3b). The overall arrangement of these SCRs suggests a strong functional link between structural domains I and II. It is also apparent that the structure of GDI can be divided into a highly conserved face that contains all the SCRs (Fig. 3b) and an opposing face that shows lower evolutionary conservation between members of the GDI and CHM families (Fig. 3c).

GCD necessary for Rab binding

The biological function common to all members of the GDI and CHM families is their ability to bind a diverse group of Rab proteins. Site-directed mutagenesis of residues within the highly conserved SCRs 1 and 3B, which together form the GCD, was used to investigate their potential role in Rab binding. Mutants of GDI were generated, expressed as recombinant His₆-tagged proteins in *E. coli* and purified to homogeneity. The location of the various mutants is shown in Fig. 4a. As negative controls for analysing the effects of substitutions within SCRs 1 and 3B, amino-acid replacements outside these regions were examined. Substitution of a serine for a lysine at residue 309 (K309S), which is found along one edge of the conserved face of domain I (Fig. 4a), had only a minor effect on binding to Rab3A ($K_d = 0.7 \mu\text{M}$) compared to wild type ($K_d = 0.45 \mu\text{M}$) (Fig. 4b). Similarly, a

mutation (K278S) on the non-conserved face at the apex of domain I (Fig. 4a) had little effect on binding ($K_d = 1.0 \mu\text{M}$). An aspartate to asparagine substitution at residue 195 of domain II (D195N), which is located in the prominent hydrophilic gorge between domains I and II (Fig. 4a), also had no detectable effect on Rab3A binding ($K_d = 0.54 \mu\text{M}$).

In contrast, mutations in SCRs 1 and 3B within the GCD led to a striking reduction in the binding of Rab3A to GDI. The tyrosine at position 39 in SCR1B, located in the loop connecting the β -strand a2 to β -strand f1 (Fig. 1c) at the apex of the GCD region (Fig. 4a), is highly conserved in both GDI and CHM family members. Substitution of this residue with valine led to a >60-fold decrease in Rab3A association ($K_d > 30 \mu\text{M}$) (Fig. 4b). The binding affinity was below the detection limits of the assay, suggesting that this substitution completely abrogates Rab3A association with GDI. The importance of residues in SCR3B was examined using five additional substitutions: E233S, R240A, T248P, T249V and M250Y (Fig. 4a). Two positions, Glu 233 and Arg 240, are completely conserved in SCR3B (GExxQGFxRxxAxxG)¹⁰ and are part of helix I (Fig. 1c) which lines the upper cleft of domain I (Fig. 4a). The other 3 residues are found in β -strand a3 (Fig. 1c) which forms the outside strand of sheet a (Fig. 4a). E233S, R240A and T248P substitutions led to a >60-fold reduction in Rab3A binding ($K_d > 30 \mu\text{M}$), whereas the more distal substitutions, T249V and M250Y led to ~24-fold ($K_d = 11 \mu\text{M}$) and 21-fold ($K_d = 9.5 \mu\text{M}$) reductions in binding (Fig. 4b). These results support the premise that the GCD represents the putative Rab binding region.

We next tested the effects of each substitution on the ability of GDI efficiently to extract the GDP-bound forms of Rab3A from membranes^{23–25}. Synaptosomes were prepared from rat brain, permeabilized to expose Rab3A bound to synaptic vesicles, and incubated with wild-type or mutant proteins (Fig. 4c). Incubation

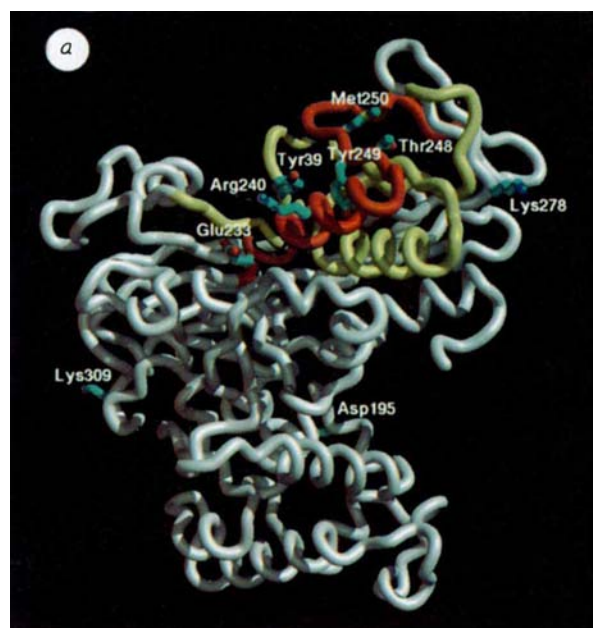
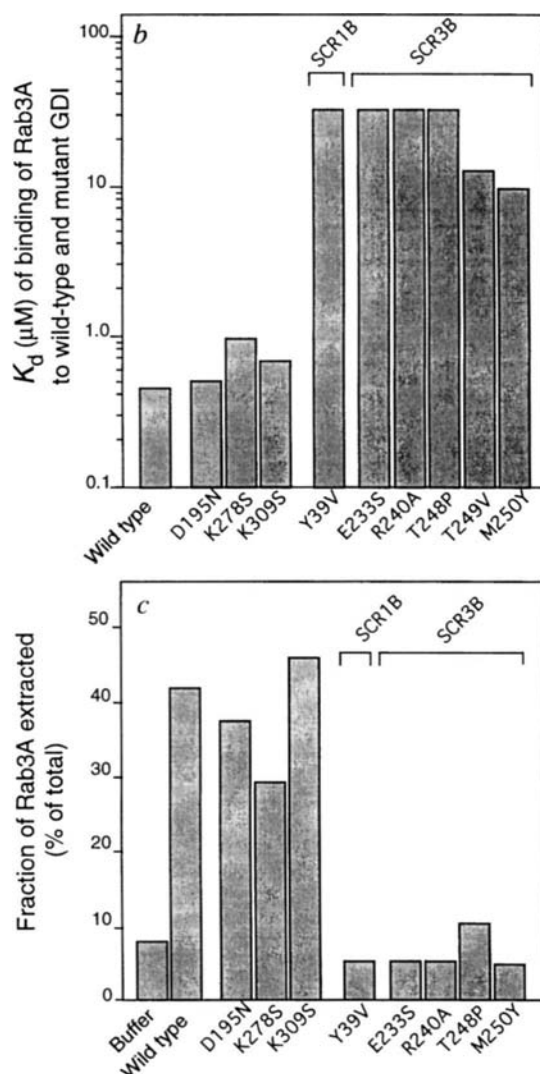


FIG. 4 *a*, Location of the mutated residues in the three-dimensional structure of GDI that were used to characterize the site of Rab binding. Side chains of individual labelled residues are shown in cyan. Glu 233, Arg 240, Thr 248, Tyr 249 and Met 250 are found in SCR3B (red). Tyr 39 is a highly conserved residue found in SCR1B (yellow). Other residues (Asp 195, Lys 278, Lys 309) are located outside the GCD subdomain (see text). *b*, Effect of site-directed mutants on the binding of GDI to Rab3 *in vitro*. The K_d values for binding of Rab3a to wild-type and mutant GDIs are shown. *c*, The effect of GDI mutants on the extraction of Rab3A from permeabilized synaptosomes. Rab3A released into the supernatant as the complex is reported as per cent of total Rab3A in the synaptosome preparation.

METHODS. Site-directed mutagenesis and purification of GDI: point mutations were introduced into bovine α -GDI by PCR site-directed mutagenesis⁴⁶. Sequence-verified mutants were subcloned into the bacterial expression vector pET11d (Novagen) containing a 5' six-histidine tag. Under induction with IPTG, mutants were expressed in *E. coli* strain BL21(DE3) and purified on NTA-agarose (Qiagen), followed by Sephacryl S-100 and Mono-Q-HR 5/5 (Pharmacia) as described⁴⁷. Preparation of prenylated Rab3A: propagation of recombinant *Autographa californica* nuclear polyhedrosis virus (AcMNPV), and growth and expression of Rab3A in Sf9 cells was carried out by the NIH National Cell Culture Center (Minneapolis, MN). Prenylated Rab3A was expressed and extracted from *Spodoptera frugiperda* (Sf9) membranes as described⁴⁸, except that purification steps were modified after NTA-agarose chromatography as follows. Eluates from the NTA-Agarose column were concentrated to 4 ml, supplemented with 0.02% Lubrol and dialysed against 100 volumes of buffer A (50 mM imidazole-HCl (pH 6.8), 75 mM NaCl, 1 mM EDTA, 0.2 DTT, 0.02% Lubrol, 10% sucrose, 5 mM $MgCl_2$, 10 μ M GDP). The sample was loaded on a Q-Sepharose Fast Flow (Pharmacia) and eluted by washing the column with buffer A. Fractions containing purified Rab3A were pooled,

of synaptosomes with wild-type GDI extracted 42% of the total Rab3A pool, as detected by western blotting, similar to its efficiency in extracting other Rab proteins^{23–25}. Consistent with its ability to form a complex with Rab3A *in vitro* (Fig. 4b), the K309S mutant efficiently extracted 47% of Rab3A from membranes (Fig. 4c). The D195N and K278S mutants, which had slightly lower affinities for Rab3A *in vitro*, were also only slightly deficient (37 and 30%, respectively) in their ability to extract Rab3A. In striking contrast, substitutions in SCRs that abolished complex formation *in vitro* (Fig. 4b) potentially inhibited the ability of GDI to extract Rab3A (Fig. 4c). The Y39V, E233S, R240A and M250Y mutants (SCR3B) were able to extract <5% of Rab3A, an amount equivalent to that obtained by washing membranes with buffer alone. We also examined whether mutations in the GCD could suppress the ability of excess GDI to inhibit transport through the early exocytic pathway²³. As expected, substitutions



concentrated 5-fold and stored in aliquots at -70°C . Determination of the binding constant for complex formation between GDI and Rab3A: dissociation constants (K_d) for complexes of prenylated Rab3A and wild-type or mutant GDI were determined by a modification of the method described in ref. 49. Under these binding conditions, Rab9 was found to have a K_d of ~ 50 nM for *E. coli* GDI, similar to the value reported previously⁴⁹. Rab3 extraction: synaptosomes were purified⁵⁰ and permeabilized as described⁵¹. Following incubation with 2 μ M (final) of wild-type GDI or the indicated mutant for 30 min on ice as described²³, membranes were pelleted and the fraction of Rab3A extracted determined using quantitative western blotting as described²³.

in SCRs 1 and 3B that prevent Rab binding completely neutralize the inhibitory properties of GDI (data not shown). Thus the ability of GDI to form a complex with purified Rab and its ability to interfere with the physiological function of Rab in vesicle traffic are closely correlated.

Implications for biological function

An apparently strong evolutionary pressure on residues within SCR1 and SCR3B has maintained the GCD structure at the apex region of domain I. Possible insight into how Rab may dock to this region comes from the structure of a Ras-related protein, Rap1A, in complex with the Ras-binding domain (RBD) of cRaf1 kinase²⁶. The outer B2 strand of a 5-stranded mixed β -sheet (strands B1–B5) in RBD interacts with an external β 2 strand of the effector domain of Rap1A to form an extended β -sheet²⁶. We have also found that residues in the effector domain of Rab3A are critical

for the binding of Rab3A to GDI (data not shown). Molecular docking studies using Ras as a model for Rab show that there is sufficient accessibility in the upper cleft of GDI to accommodate Rab in an extended β -sheet structure, although the exact mode of interaction will have to await structure determination of the Rab-GDI complex. Given that the conformation of the effector domain of Ras superfamily GTPases is particularly sensitive to the GDP- or GTP-bound state of the protein, binding through the effector domain may explain why GDI can recognize only GDP-bound forms of Rab. The other two sequence-conserved regions, SCR2 and SCR3A, which form the lower half of the conserved face of GDI (Fig. 3b, c), represent candidate regions for interaction with receptors mediating Rab binding to membranes.

The structure of GDI provides important insights into the structural organization of CHM. Sequence alignment of the α - and β -isoforms of GDI with CHM and CHML shows regions of identity which not only include SCRs, but many other residues that are distributed throughout the entire primary sequence (see Supplementary Information). It is likely that the structural architecture of members of the CHM family is similar to that of α -GDI. This conclusion would be consistent with the almost identical function of CHM and GDI in the delivery and extraction of Rab from membranes¹⁴ and the possible requirement for the effector domain in protein prenylation²⁷. The major difference between GDI and CHM family members is the large insert (138 amino acids in human/mouse CHM/CHML) between SCRs 1 and 2 which replaces the 20–30-residue insert (β -strand strands f1, f2 and helix B (Fig. 2b, green)) connecting SCRs 1 and 2 in GDI (Fig. 3a). This insert may provide the basis for unique CHM functions that involve interaction with the catalytic components of Rab geranylgeranyl transferase II and/or binding to both prenylated and non-prenylated forms of Rab.

The relationship between CHM structure and disease is now

evident. Point mutations in Danish and Swedish CHM genes leading to choroideaemia give rise to premature stop codons⁹ (Fig. 3a, arrows). Immortalized lymphoblasts prepared from patients deficient in CHM mRNA have a marked reduction in component A (REP1) activity of geranylgeranyl transferase II (ref. 8). Deletion of only 70 C-terminal residues leads to loss of function⁹ and would be expected to include helix N of domain I. This change would cause a major perturbation in the integrity of domain I and hence provide a structural basis for changes that lead to the diseased state.

Outlook

To our knowledge, this is the first crystallization and structural characterization of an intact regulatory protein interacting with the Ras superfamily of GTPases. The function of GDI in membrane vesicular transport is not unlike that of $\beta\gamma$ subunits of heterotrimeric G proteins involved in signal transduction. The G-protein $\alpha\beta\gamma$ complex, whose structure has been determined^{28,29}, serves as a membrane-associated reservoir for the activation and release of the GTPase α -subunit from the $\beta\gamma$ subunits in response to a wide range of receptor–ligand complexes. In an analogous fashion, the Rab–GDI complex provides a cytosolic reservoir for recruitment and activation of Rab. Rab is likely to be the specificity subunit that contributes selectivity to the interaction of GDI complexes with specific organelles of the endocytic and exocytic pathway. Given the pivotal role of GDI in Rab function, its structure provides a foundation to explore the mechanisms by which membrane receptors and ligands regulate Rab delivery to control vesicle targeting and fusion. The structural similarity between flavoproteins and GDI raises the possibility that FAD may be a cofactor for an unsuspected GDI hydroxylase or oxidase activity which may participate in regulating membrane transport. □

Received 17 January; accepted 21 March 1996.

- Pfeffer, S. R., Dirac-Sveinsson, B. & Soldati, T. *J. biol. Chem.* **270**, 17057–17059 (1995).
- Nuoffer, C., Balch, W. E. *A. Rev. Biochem.* **33**, 949–990 (1994).
- Sasaki, T. *et al. J. biol. Chem.* **265**, 2333–2337 (1990).
- Ullrich, O. *et al. J. biol. Chem.* **268**, 18143–18150 (1993).
- Anders, D. A. *et al. Cell* **73**, 1091–1099 (1993).
- Creemers, F. P. M. *et al. Hum. molec. Genet.* **1**, 71–75 (1992).
- Seabra, M. C., Brown, M. S., Slaught, C. A., Sudhof, T. C. & Goldstein, J. L. *Cell* **70**, 1049–1057 (1992).
- Seabra, M. C., Brown, M. S. & Goldstein, J. L. *Science* **259**, 377–381 (1993).
- van Bokhoven, H. *et al. Hum. molec. Genet.* **3**, 1041–1046 (1994).
- Waldherr, M., Ragnini, A., Schweyen, R. J. & Boguski, M. S. *Nature Genet.* **3**, 193–194 (1993).
- Creemers, F. P. M., Armstrong, S. A., Seabra, M. C., Brown, M. S. & Goldstein, J. L. *J. biol. Chem.* **269**, 2111–2117 (1994).
- Jiang, Y. & Ferro-Novick, S. *Proc. natn. Acad. Sci. U.S.A.* **91**, 4377–4381 (1994).
- Fujimura, K., Tanaka, K., Nakano, A. & Toh-e, A. *J. biol. Chem.* **269**, 9205–9212 (1994).
- Alexandrov, K., Horiuchi, H., Steele-Mortimer, O., Seabra, M. C. & Zerial, M. *EMBO J.* **13**, 5262–5273 (1994).
- Entsch, B. & van Berk, W. J. H. *FASEB J.* **9**, 476–483 (1995).
- Hecht, J. J., Kalisz, H. M., Hendle, J., Schmid, R. D. & Schomburg, D. *J. molec. Biol.* **229**, 159–172 (1993).
- Li, J., Vrelink, A., Brick, P. & Blow, D. M. *Biochemistry* **32**, 11507–11515 (1993).
- Holm, L. & Sander, C. *J. molec. Biol.* **233**, 123–138 (1993).
- Schreuder, H. A. *et al. J. molec. Biol.* **208**, 679–696 (1989).
- Wierenga, R. K., de Jong, R., Kalk, K. H., Hol, W. G. J. & Drenth, J. *J. molec. Biol.* **131**, 55–73 (1979).
- Wierenga, R. K., Terpstra, P. & Hol, W. G. J. *J. molec. Biol.* **187**, 101–107 (1986).
- Stott, K., Saito, K., Thiele, D. J. & Massey, V. *J. biol. Chem.* **268**, 6097–6106 (1993).
- Peter, F., Nuoffer, C., Pind, S. N. & Balch, W. E. *J. Cell Biol.* **126**, 1393–1406 (1994).
- Soldati, T., Riederer, M. A. & Pfeffer, S. R. *Molec. Biol. Cell* **4**, 425–434 (1993).
- Ullrich, O., Horiuchi, H., Bucci, C. & Zerial, M. *Nature* **368**, 157–160 (1994).
- Nassar, N. *et al. Nature* **375**, 554–560 (1995).
- Wilson, A. L. & Maltese, W. A. *J. biol. Chem.* **268**, 14561–14564 (1993).
- Lambright, D. G. *et al. Nature* **379**, 311–319 (1996).
- Wall, M. A. *et al. Cell* **83**, 1047–1058 (1995).
- Schalk, I. J., Stura, E. A., Matteson, J., Wilson, I. A. & Balch, W. E. *J. molec. Biol.* **244**, 469–473 (1994).
- Stura, E. A., Satterthwaite, A. C., Calvo, J. C., Kaslow, D. C. & Wilson, I. A. *Acta crystallogr.* **50**, 448–455 (1994).
- Matthews, B. W. *J. molec. Biol.* **33**, 491–497 (1968).
- Hamson, S. C. *J. appl. Crystallogr.* **1**, 84–90 (1968).
- Howard, A. J. *et al. J. appl. Crystallogr.* **20**, 383–387 (1987).
- McRee, D. E. *J. molec. Graph.* **10**, 44–47 (1992).
- Furey, W. *Am. crystallogr. Assoc. 40th Anniv. Meeting* (New Orleans, LA, 1990).
- Wang, B. C. *Meth. Enzym.* **115**, 90–112 (1985).
- Brünger, A. T. thesis, Yale Univ. (1992).
- McRee, D. E. *Practical Protein Crystallography* 1–225 (Academic, San Diego, 1993).
- Brünger, A. T. *Nature* **355**, 472–474 (1992).
- Laskowski, R. A., MacArthur, M. W., Moss, S. D. & Thornton, J. M. *J. appl. Crystallogr.* **26**, 283–291 (1993).
- Kraulis, P. J. *J. appl. Crystallogr.* **24**, 946–950 (1991).
- Kabsch, W. & Sander, S. *Biopolymers* **22**, 2577–2637 (1993).
- van den Hurk, J. A. J. M. *et al. Am. J. hum. Genet.* **50**, 1195–1202 (1992).
- Ferrin, T. E., Huang, C. C., Jarvis, L. E. & Langridge, R. J. *J. molec. Graph.* **6**, 13–27 (1988).
- Dascher, C. & Balch, W. E. *J. biol. Chem.* **269**, 1437–1448 (1994).
- Nuoffer, C. N., Davidson, H. W., Matteson, J., Meinke, J. & Balch, W. E. *J. Cell Biol.* **125**, 225–237 (1994).
- Nuoffer, C., Peter, F. & Balch, W. E. *Meth. Enzym.* **257**, 3–8 (1995).
- Shapiro, A. D. & Pfeffer, S. R. *J. biol. Chem.* **270**, 11085–11090 (1995).
- Fisher von Mollard, G., Sudhof, T. C. & Jahn, R. *Nature* **349**, 79–81 (1991).
- Dekker, L. V., De Graan, P. N. E., Pijnappel, P., Oestreicher, A. B. & Gispen, W. H. *J. Neurochem.* **54**, 205–210 (1991).
- Jones, T. *et al. Acta crystallogr.* **A47**, 110–119 (1991).

ACKNOWLEDGEMENTS. I.S. and K.Z. contributed equally to this work. This work was supported by grants from the NIH (to W.E.B. and I.A.W.), the Tobacco-related Disease Research Program of California (to W.E.B.) and the Lucille B. Markey Charitable Trust. A.T. is a recipient of fellowships from the Fonds de la Recherche en Santé du Québec and the MRC of Canada. We thank the National Cell Culture Center, NIH, for propagation of recombinant *Autographa californica* nuclear polyhedrosis virus, and for growth and expression of Rab3A in Sf9 cells and S. Pfeffer (Stanford University) for Rab9.

CORRESPONDENCE and requests for materials should be addressed to I.A.W. or W.E.B. Coordinates will be deposited in the Brookhaven databank.

SUPPLEMENTARY INFORMATION. Available on Nature's World-Wide Web site <http://www.nature.com>; requests may also be addressed to Mary Sheehan at the London editorial office of Nature.

Dependence of the spectral evolution of γ -ray bursts on their photon fluence

Edison Liang* & Vincent Kargatis†

* Department of Space Physics and Astronomy, Rice University, Houston, Texas 77005, USA

† Hughes STX, Greenbelt, Maryland 20771, USA

THE nature of γ -ray bursts remains one of the great mysteries of modern astronomy^{1,2}. Although the spatial distribution of these high-energy sources is tightly constrained (they are distributed isotropically across the sky³), the same cannot be said of their spectral and temporal properties, for which few clear trends have been identified^{1,2}. But it is from such properties that important clues about the physical mechanisms underlying the bursts are likely to be derived^{4,5}. Here we show that for a sample of bursts consisting of well resolved, isolated energy pulses, and for which the spectra are sufficiently clean to allow their temporal evolution to be modelled, the photon energy at which the power output is maximum for each pulse decreases exponentially with photon fluence (that is, running time integral of photon flux). We find that for many multi-pulse bursts, the exponential decay constant is invariant from pulse to pulse. This property favours models in which the pulses arise from a regenerative source, rather than as a consequence of a single catastrophic event.

We model the evolution of a γ -ray burst (GRB) observed by the Burst and Transient Source Experiment (BATSE) by fitting the time-resolved count spectra via χ^2 -minimization with the empirical Band *et al.*⁶ model consisting of two power laws joined smoothly by an exponential. It has three parameters: two power-law slopes and the exponential constant, which is proportion to the peak power energy, E_{pk} (defined as the maximum of νF_ν , where F_ν is specific energy flux ($\text{erg cm}^{-2} \text{s}^{-1} \text{keV}^{-1}$) and ν is photon energy in keV), plus overall normalization. To focus on the evolution of E_{pk} we fix the two slopes throughout the burst by fitting the Band *et al.* model⁶ to the time-averaged spectrum. Figure 1 shows the result of such spectral modelling for a typical BATSE burst. Based on the experience of Ford *et al.*⁷ we choose time bin widths such that the signal-to-noise ratio in the 60–200 keV band is at least 30–40, to provide enough photons for reliable spectral fits. Hence time bins are wide when the flux is low and vice versa. The trade-off between time bin width and signal-to-noise ratio is often optimized iteratively. The binned time history of E_{pk} is compared to those of the photon and energy flux obtained by deconvolving the count data with the best-fit model. We see in Fig. 1 that the first pulse evolves from ‘hard-to-soft’^{7–9} while in the second pulse E_{pk} ‘tracks’ the flux^{9–11} within our limited time resolution. Both behaviours are common among GRB pulses^{4–8}.

In Fig. 2 inset we plot E_{pk} versus photon flux for the same burst, showing that there is no simple relation between E_{pk} and flux. But when we plot E_{pk} as a function of photon fluence (running time-integral of the flux) in Fig. 2 main figure, we see that during the decay E_{pk} falls exponentially as a function of photon fluence. We have analysed 34 bright BATSE bursts with long smooth pulse decays that are most suitable for time-resolved spectral modelling.

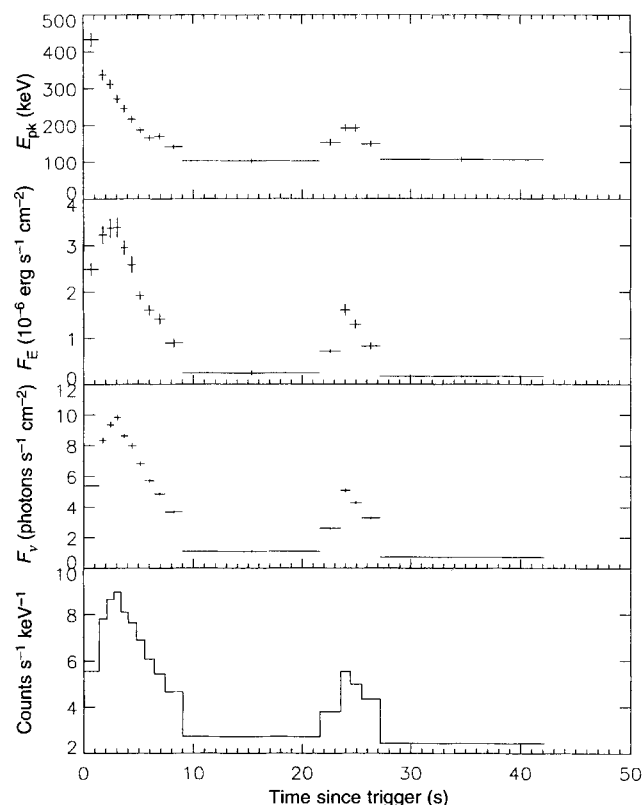


FIG. 1 Time histories of BATSE trigger no. 0973 (GRB0973). The four panels are plots of the time profiles (top to bottom) of E_{pk} , energy flux, photon flux and channel-averaged count rate.

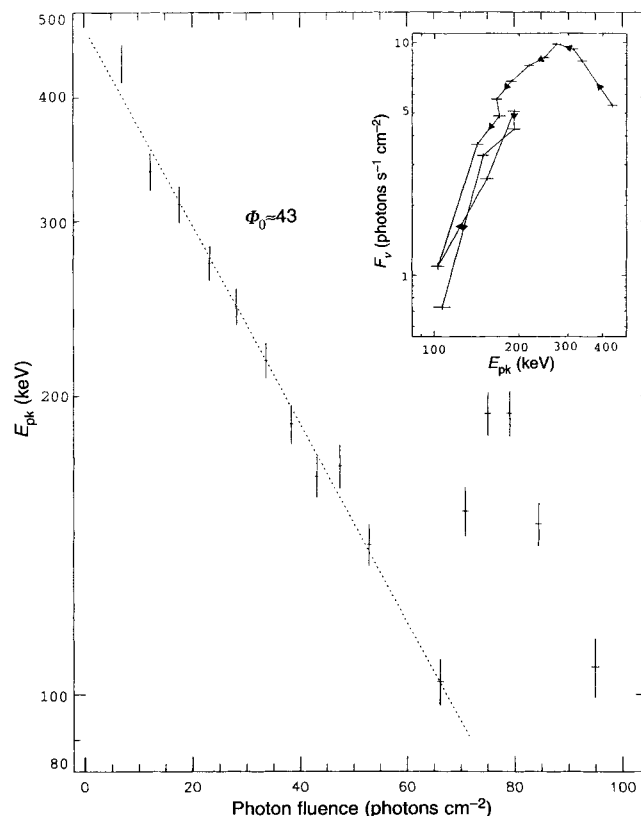


FIG. 2 inset, E_{pk} versus photon flux for GRB0973 showing no obvious relation between the two variables. Main figure, E_{pk} versus photon fluence for the same burst. Dotted line is the best exponential fit to the first pulse decay (compare Fig. 1). Label next to the dotted line is the best-fit decay constant Φ_0 . Error bars are 1σ . Fluences are evaluated at the end of each time bin. We do not fit the second pulse decay as it has only three time bins (see text).

We find that the exponential decay of E_{pk} versus photon fluence, although not universal, is very common. These 34 bursts are selected from ~ 400 bursts detected up to September 1994, and stored in the BATSE Spectroscopy Detector database. Typically we prefer that the brightest pulse in the burst be long and smooth enough to provide at least seven useful time bins, or that the overall burst contains at least 10 useful time bins and two or more pulses for comparative studies. Though 34 bursts constitute a small subset of all BATSE bursts, we suspect that this behaviour is common among other bursts, but it may be difficult to discern in weaker, shorter or more complex bursts (another ~ 100 bursts are currently being analysed but the above selection effect probably will only be alleviated when we have useful higher-time-resolution data).

For the above analysis we define a 'pulse decay' as a stretch of time bins starting from a local maximum of E_{pk} coincident with or following a rise in flux (that is, increases by $> 2\sigma$) or before the flux vanishes. The dashed lines in Figs. 2 and 3 are the best χ^2 -fits to those bins within a clearly defined pulse decay. Table 1 summarizes the statistics of 37 pulses (from 34 GRBs) whose E_{pk} decays last four or more time bins (pulses with three bins or less are statistically too marginal to quantify). Only two such pulses are inconsistent with exponential decay as a function of fluence; the poor fit in both these cases is due to deviant data in one time bin (for example, the second pulse of GRB3138 in Fig. 3).

Figure 3 shows examples of multi-pulse bursts in which the exponential decay constants of the subsequent pulses are, within statistical errors, consistent with that of the leading pulse. This suggests that the physical parameters determining the decay constant may be invariant from pulse to pulse. Of the 12 GRBs that we have analysed with multiple pulses lasting four time bins or longer, seven are consistent with an invariant decay constant (difference $< 1\sigma$), one is marginally consistent ($< 3\sigma$), while four are not ($> 3\sigma$). Hence the invariance of the decay constant may be the signature of a subclass of GRBs.

The above pulse-decay data implies that the characteristic GRB photon energy $\langle \nu \rangle$ as represented by E_{pk} evolves as $\langle \nu \rangle(t) = \langle \nu \rangle_0 \exp(-\Phi(t)/\Phi_0)$, where $\langle \nu \rangle_0$ is the peak value of $\langle \nu \rangle$ within a pulse, $\Phi(t)$ is the photon number fluence integrated from the time of $\langle \nu \rangle_0$ and Φ_0 is the decay constant. Differentiating this gives

$$-d\langle \nu \rangle/dt = \langle \nu \rangle F_N(t)/\Phi_0 \approx F_E(t)/\Phi_0 = L(t)/\Omega d^2 \Phi_0 \quad (1)$$

where F_N is photon number flux (photons $\text{cm}^{-2} \text{s}^{-1}$), F_E is energy flux ($\text{erg cm}^{-2} \text{s}^{-1}$), d is the distance to the burster, Ω is the beam angle ($= 4\pi$ for isotropic emission) and L is total luminosity (erg s^{-1}). The second equality of equation (1) assumes that $\langle \nu \rangle$ is approximately equal to the average photon energy in the BATSE band (which we have checked to be generally true) and the last equality assumes that Ω remains constant during the pulse. Hence the rate of decrease of the characteristic photon energy is proportional to the instantaneous γ -ray luminosity of the source. This empirical result holds for both 'hard-to-soft'⁷⁻⁹ and 'tracking'^{10,11} pulses. Figure 1 shows why such a result was hard to uncover in the time profiles: with the very limited time resolution available to perform the spectral modelling, it is practically impossible to compute accurately the time derivative of E_{pk} . The correlation becomes evident only when we integrate equation (1). (The linear relation between E_{pk} and energy fluence, though visible in many bursts, is statistically harder to establish than the E_{pk} versus photon-fluence relation because the deconvolved energy flux

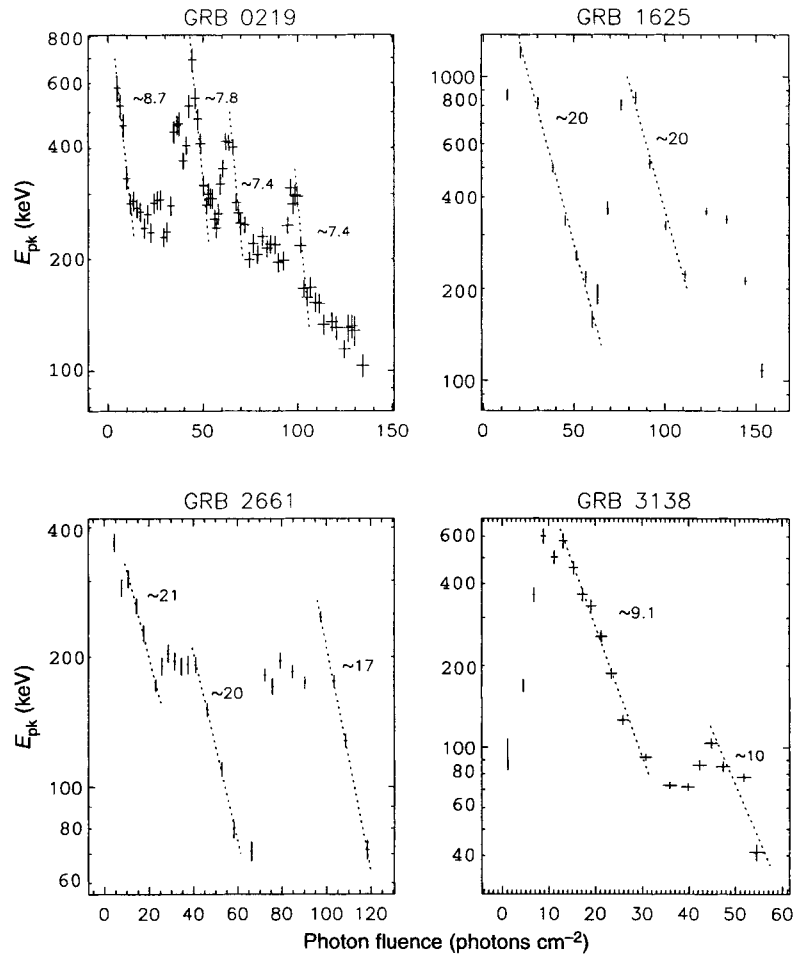


FIG. 3 E_{pk} versus photon fluence for GRB0219, GRB1625, GRB2661 and GRB3138, showing multiple pulses with almost invariant decay constant Φ_0 . Only pulse decays lasting four or more times bins are analysed.

typically has much larger errors than the photon flux, due to the contributions of the poorly measured high-energy channels.)

One interesting interpretation of equation (1) is that of a confined plasma with a fixed number of particles N cooling via γ -radiation. (We emphasize that this interpretation is not unique and that we are exploring alternatives.) If the average particle energy is $\langle E \rangle$, then by energy conservation

$$L = -N d\langle E \rangle/dt \quad (2)$$

which is consistent with equation (1) if $\langle \nu \rangle \propto \langle E \rangle$. This is indeed

TABLE 1 Statistics of E_{pk} versus photon fluence decay profiles from 34 BATSE GRBs

Total number of pulses analysed with decay lasting four time bins or longer	37
Number of pulses consistent with exponential decay*	25
Number of pulses marginally consistent with exponential decay†	10
Number of pulses not consistent with exponential decay‡	2

* Consistency is defined as Pearson correlation coefficient¹⁷ $R \geq 0.98$ and χ^2 -significance¹⁷ $Q > 10^{-3}$ when fitted with a straight line in a log-linear plot (Figs 2, 3). Due to the small number of time bins in most cases it is impractical to define the goodness of fit more rigorously.

† Marginal consistency is defined as $0.98 > R > 0.95$ and $Q > 10^{-3}$ or $Q < 10^{-3}$ if $R \geq 0.98$.

‡ Inconsistency is defined as $R \leq 0.95$ or $Q < 10^{-3}$ if $0.98 > R > 0.95$.

the case if the escaping photons share the average energy of the emitting particles, such as in multiple Compton scattering^{12–14}. To constrain further the radiation mechanism¹² equation (1) must be combined with detailed time-domain information and a model for the origin of the spectral break^{13,14}.

Combining equations (1) and (2) we have

$$\Phi_0 = Nf/\Omega d^2 \quad (3)$$

where $f = \langle E \rangle / \langle v \rangle$ is a constant of the order of unity. For those bursts in which Φ_0 is invariant for repeating pulses as in Fig. 3, equation (3) implies that those pulses retain the same number of

emitting particles, hinting that they originate from the same burst site and that the plasma is confined. This would favour astrophysical models in which the different pulses are produced by a regenerative source rather than a single catastrophic event¹⁵. Though the above interpretation is independent of the specific radiation mechanism, it suggests that the spectral softening and pulse-decay timescale¹⁶ are controlled by the radiative cooling process and not by the energy supply. Independent of its interpretation, equation (1), if confirmed by data with higher time-resolution, should greatly influence astrophysical modelling of GRBs. □

Received 6 December 1995; accepted 22 March 1996.

1. Fishman, G. J. & Meegan, C. A. *Rev. Astr. Astrophys.* **33**, 415–458 (1995).
2. Hurley, K. in *Gamma Ray Bursts* (eds Paciesas, W. & Fishman, G.) 3–12 (Conf. Proc. 265, Am. Inst. Phys. New York, 1992).
3. Meegan, C. A. et al. *Astrophys. J.* (submitted).
4. Schaefer, B. et al. *Astrophys. J.* **393**, L51–L54 (1992).
5. Liang, E. in *Gamma Ray Bursts* (eds Paciesas, W. & Fishman, G.) 246–251 (Conf. Proc. 265, Am. Inst. Phys. New York, 1992).
6. Band, D. et al. *Astrophys. J.* **413**, 281–292 (1993).
7. Ford, L. et al. *Astrophys. J.* **439**, 307–321 (1995).
8. Norris, J. et al. *Astrophys. J.* **301**, 213–219 (1986).
9. Ford, L., Band, D., Matteson, J., Teegarden, B. & Paciesas, W. in *Gamma Ray Bursts* (eds Fishman, G. et al.) 298–302 (Conf. Proc. 307, Am. Inst. Phys. New York, 1994).
10. Kargatis, V. et al. in *Compton Gamma-Ray Observatory* (eds Friedlander, M. et al.) 907–911 (Conf. Proc. 280, Am. Inst. Phys. New York, 1993).

11. Bhat, P. N. et al. in *Compton Gamma-Ray Observatory* (eds Friedlander, M. et al.) 912–916 (Conf. Proc. 280, Am. Inst. Phys. New York, 1993).
12. Rybicki, G. & Lightman, A. *Radiative Processes in Astrophysics* (Wiley, New York, 1979).
13. See Lamb, D. Q. in *Nuclear Spectroscopy of Astrophysical Sources* (eds Gehrels et al.) 265–284 (Conf. Proc. 170, AIP, 1987).
14. Harding, A. K. *Phys. Rep.* **206**, 327–394 (1991).
15. Meszaros, P. & Rees, M. J. *Astrophys. J.* **405**, 278–284 (1993).
16. Fenimore, E., in 't Zand, J., Norris, J., Bonnell, J. & Nemiroff, R. *Astrophys. J.* (in the press).
17. Press, W., Teukolsky, S., Vetterling, W. & Flannery, B. *Numerical Recipes* 609–699 (Cambridge, UK, 1992).

ACKNOWLEDGEMENTS. The authors thank the BATSE team for help in this work. This work was supported by NASA; V.K. thanks NASA for a graduate student research program fellowship.

CORRESPONDENCE should be addressed to E.L.

Formation of asteroid satellites and doublet craters by planetary tidal forces

William F. Bottke Jr* & H. Jay Melosh†

* Division of Geological and Planetary Sciences, California Institute of Technology, M/S 170-25, Pasadena, California 91125, USA

† Lunar and Planetary Laboratory, University of Arizona, Tucson, Arizona 85721-0092, USA

APPROXIMATELY ten per cent of the impact structures on the Earth and Venus are doublets^{1,2}—pairs of craters formed by the near-simultaneous impact of asteroids of comparable size. It has been suggested that these doublet craters form from asteroid fragments dispersed by aerodynamic forces during atmospheric entry^{1,3}, or from asteroids that were tidally disrupted by gravitational forces shortly before impact^{4–6}. But to form a doublet, the progenitors of the craters must have been well separated before final impact¹, which poses problems for both mechanisms. Here we argue that a hitherto undetected population of well separated binary asteroids can explain the occurrence of doublet craters. By modelling asteroids as weak, gravitationally bound aggregates ('rubble piles'), we show that the tidal forces experienced during close encounters with the Earth can generate binary asteroids, in a process similar to that which fragmented the comet Shoemaker–Levy 9 (ref. 7) as it passed by Jupiter. Although the resulting binary asteroids may eventually separate or coalesce before colliding with a planet, repeated close encounters with the Earth maintain a steady-state population that is sufficiently large to explain the observed number of doublet craters.

At least 3 of the 28 largest impact structures on Earth with diameters greater than 20 km have a nearby companion crater sharing the same formation age¹. Among the other terrestrial planets, only Venus has been surveyed quantitatively for doublet craters². Crater pairs on Venus are identified by their close association and/or by morphology indicating simultaneous formation (that is, shared ejecta blankets and crater walls, with no overlapping features). Results from these surveys show that Venus's observed doublet-crater population fraction is smaller

than Earth's (only 2–3%), though one must take account of Venus's thick atmosphere which shields its surface from small impactors. If Earth's atmosphere were as thick as Venus's, only one of its three doublet craters would have been formed, making its proportion of doublets essentially the same as Venus's. In addition, by surveying the fraction of doublet 'splotches', radar-dark features created by projectiles so small (< 200 m) that they catastrophically disrupt in Venus's atmosphere, we find a representation of Venus's impactor population (~ 14%) which is consistent with Earth's fraction of doublet craters (~ 10%).

We suggest that asteroid satellites, generated as a by-product of close encounters between asteroids and planets, can produce the fraction of doublet craters observed on Earth and Venus. To quantify this hypothesis numerically, we model close encounters between loosely bound contact-binary asteroids and the Earth using an adaptive fifth-order Runge–Kutta numerical integrator^{1,8}. For each test case, we use a Monte Carlo method, starting with 10,000 contact-binary asteroids initially far from the Earth with randomly chosen initial orientations, a selected encounter velocity V_{∞} , and a selected close approach distance d (Fig. 1). These objects are assumed to be initially in mutual contact but co-orbiting each other, and so have periods of the order of 4 hours. Our results show that contact-binaries can be tidally perturbed into co-orbiting asteroids, though the separation distance between the components is almost always too small to produce a doublet crater (that is, in most cases, the separation distance is only a few times the mean diameter of the original asteroid).

To first order, our results agree with the limited outcome statistics of more sophisticated (and time consuming) N -body codes treating encounters between rubble-pile asteroids and the Earth (ref. 9 and D. R. Richardson, personal communication). Rubble-pile asteroids are defined as a collection of gravitationally self-bound components ranging in size from micrometres up to 100 m or kilometre-sized fragments (that is, they are not dust piles). Similar to contact-binaries, rubble-piles undergo 'mass-stripping', implying that only small (~ 100-m) individual fragments are created or ejected, or 'tidal fission', implying similar-sized components were created, during close Earth encounters if they have one (or more) of the following characteristics: (1) a fast prograde rotation rate (near the critical breakup limit), (2) an elongated shape (mass is more readily shed from the ends of the asteroid), (3) periape distance close to Earth, (4) low

encounter velocities, and/or (5) low bulk density. Moreover, if more than two fragments are stripped from a rubble-pile, the multiple-fragment system evolves into a binary system where the extra fragments either collide with a bound fragment or escape¹⁰. Thus, because fragment clusters should evolve similarly to star clusters, where the most stable end-states are binary systems, both models yield similar orbiting outcome results.

Once the asteroid's components are orbiting one another, they become susceptible to further small perturbations during repeated distant Earth encounters and mutual tidal forces which modify their mutual separation distance. To model these effects, we combined our numerical model of planetary encounters (described above) with another Monte Carlo code¹¹ that computes the frequency and characteristics of repeated multiple encounters with Earth along with mutual tidal effects between the components. The Monte Carlo code uses a probability distribution for asteroid velocities at encounter based on actual orbits of Earth-crossing asteroids¹². Our results show that if an asteroid satellite is formed during a close approach with Earth, distant perturbations often substantially increase the separation distance between the components in a random-walk fashion; components can also escape or collide with one another. Tidal forces between the components can also increase (or decrease) the separation distance, though their effect decreases substantially when the components are far apart.

Even though the scheme we have modelled produces a large number of well separated binary asteroids, we find that most of these binary asteroids do not survive to strike the Earth; close approaches with Earth, which are more probable than impact

encounters, nearly always cause the binary's components to escape one another. Thus, if contact-binary asteroids only produce a single satellite from planetary tidal forces, few (if any) doublet craters would be formed. Even an initial population of binary asteroids from the main asteroid belt would not enhance the observed fraction of doublet craters on Earth because most secondary components would be lost before impact.

However, rubble-pile asteroids have the potential to produce an asteroid satellite each time they have a close encounter with the Earth, conceivably replacing any satellite lost during a previous (or the same) encounter. As long as these rubble piles maintain sufficient mass, mass-stripping of small components off the primary or even tidal fission may occur many times over their lifetime. But these secondary components, if small enough, would probably be unable to produce a doublet crater on Venus (or even the Earth) because atmospheric screening would prevent them from striking the surface. This screening process prevents the formation of craters caused by satellites smaller than ~ 100 m diameter on Earth or 1 km on Venus.

Using this concept of continuous rebirth of secondary satellites during close encounters, we can determine the steady-state distribution of binary asteroids in the Earth-crossing region (Fig. 2). We find that a population of weakly bound single asteroids evolves into a population where over half ($\sim 57\%$) the objects are binary asteroids, some separated by large distances which are only limited by solar tides (solar tides cause the dispersion of components with radii of 0.5 and 1.0 km separated by more than ~ 60

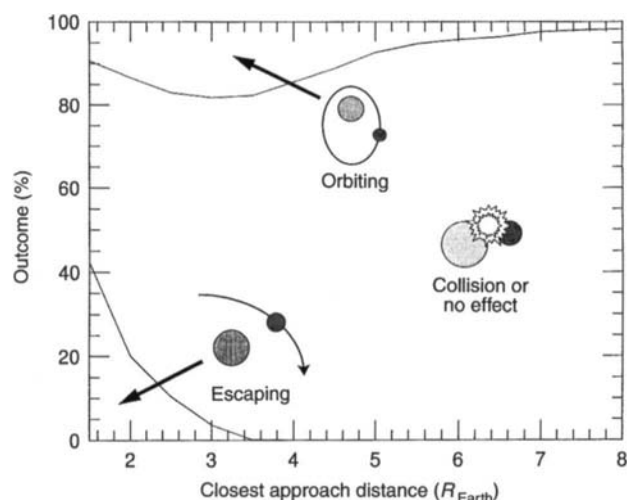


FIG. 1 Post-encounter statistics for spherical contact-binary asteroids (0.5 and 1 km in radius with a rotation period of 3.55 hours and a density of $2,600 \text{ kg m}^{-3}$) encountering the Earth at velocity $V_{\infty} = 12 \text{ km s}^{-1}$ over various closest-approach distances (d between 1.5 and 8.0 Earth radii, in increments of 0.5 Earth radii). Over 10,000 random initial orientations were used for each choice of V_{∞} and d . The motion starts and finishes at a distance of 60 Earth radii. (The V_{∞} value of 12 km s^{-1} represents the average encounter velocity between the Earth-crossing asteroids and the Earth¹².) The three encounter outcomes, (1) components escaping one another, (2) components colliding with one another or no effect from planetary tides, and (3) components orbiting one another, were tabulated and plotted (against d) as a percentage, where all the outcomes added together equal 100%. As d increases, the fraction of escape outcomes decreases, demonstrating that planetary tides weaken as the asteroids move away from the planet (the planetary tides decrease as $1/d^3$). The fraction of orbiting outcomes increases and then decreases as d increases, showing a maximum near $d = 3.0$ – 3.5 Earth radii where escape outcomes become extremely unlikely. The median semimajor axis between the components for each set of orbiting outcomes at a given distance d is 3.4 km or smaller while the median eccentricities are 0.50 or smaller, making the separation distance between the components too small to produce a doublet crater except under exceptional conditions.

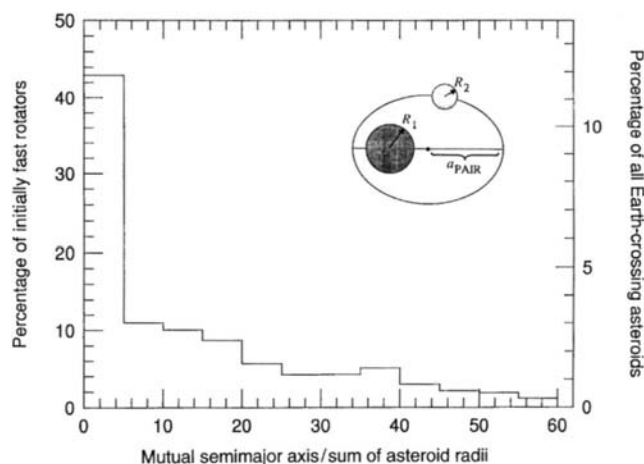
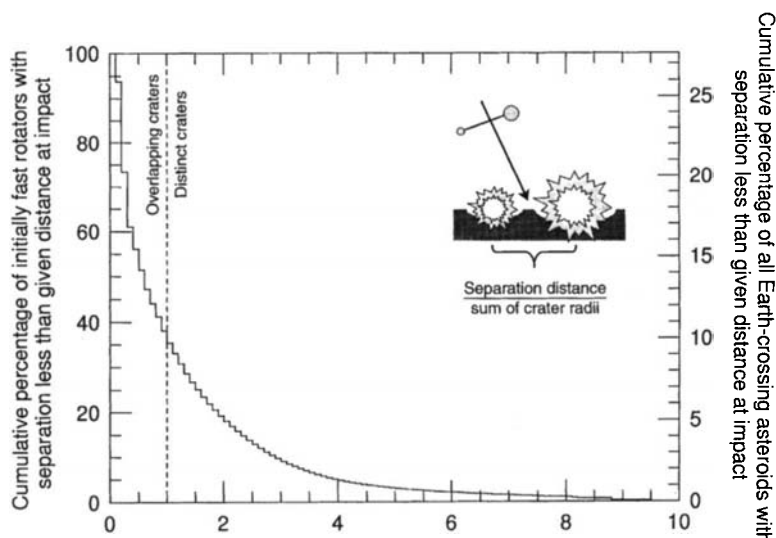


FIG. 2 The steady-state distribution of co-orbiting asteroids in the Earth-crossing region is found by allowing spherical contact-binary asteroids (0.5 and 1 km in radius with a rotation period of 3.55 hours and a density of $2,600 \text{ kg m}^{-3}$) to evolve over multiple encounters with the Earth using a probability distribution of relative encounter velocities V_{∞} (2.0 – 38 km s^{-1} in increments of 4 km s^{-1})¹². At least 90 bodies were run in the Monte Carlo code for each velocity increment. Each asteroid can evolve for as long as 100 Myr, the typical lifetime of Earth-crossing asteroids (ECAs) against planetary collision¹², though many do not survive that long. If a close Earth encounter produces a co-orbiting asteroid, distant Earth encounters then modify the newly formed binary asteroid's mutual semimajor axis, eccentricity and inclination. In the interim between Earth encounters, mutual tidal forces between the components often modify the binary's semimajor axis, eccentricity and rotation period of the larger component by exchanging rotational and orbital angular momentum¹¹. The ordinate on the left-hand side of the plot shows the percentage of objects starting with this 3.55-hour rotation period that evolve into binary asteroids; the abscissa shows their mutual semimajor axis (a_{PAIR}) in units of the sum of the radii of the components ($R_1 + R_2$). Our results show that our starting asteroid population evolves into a population where over half are binaries, some separated by large distances. The mutual eccentricities of those binaries tend to be small (most orbit each other on nearly circular orbits). The ordinate on the right-hand side of the plot shows the percentage of the kilometre-sized ECAs that should have satellites ($\sim 15\%$). We obtain this percentage by scaling our results by the actual fraction of ECAs with short rotation periods ($\sim 28\%$ with rotation periods < 5.5 hours; see text for details).

FIG. 3 The fraction of Earth-crossing asteroids (ECAs) striking Earth that produce doublet craters, calculated by putting the results of Fig. 2 into a model which numerically integrates impact encounters between binary asteroids and the Earth¹. Encounter velocities V_{∞} and other asteroid parameters are the same as described in Fig. 2. The abscissa shows the separation distance between binary asteroid components at impact divided by the sum of the crater radii found using crater scaling-law results^{16,17}; a value greater than one means that the binary creates a doublet crater, and a value less than one means that the craters overlap one another. The ordinate on the left-hand side of the plot shows the cumulative percentage of objects with 3.55-hour rotation periods that strike Earth at a given separation distance or smaller. The ordinate on the right-hand side of the plot shows the cumulative number of kilometre-sized ECAs that strike Earth at a given separation distance or smaller. Our results show that $\sim 10\%$ of the ECAs striking the Earth produce doublet craters.



mean diameters at 1.0 AU). But these results only take into account rapidly rotating asteroids near the theoretical breakup limit; many Earth-crossing asteroids (ECAs) are slow rotators with rotation periods longer than ~ 6 hours, making them difficult to pull apart by Earth's tidal forces. To find the fraction with short rotation periods, we modelled the rotation-period distribution of the ECAs as a maxwellian distribution with a mean period of 6 hours, scaling the distribution for extremely slow rotators like 4179 Toutatis ($\sim 20\%$ of the ECA population) and the paucity of rotation periods lower than 3.55 hours (ref. 13 and A. W. Harris, personal communication). It is even possible that tidal forces themselves may be the source of rapid rotation: tidal torques imparted to rubble-pile asteroids frequently increase their rotation rate, such that multiple planetary encounters may serve to spin up the body to near-disruption (A. W. Harris, personal communication). Thus, the same tidal process that disrupts critically spinning bodies may be responsible for bringing bodies to near-critical spin rates. A thorough analysis of this effect is beyond the scope of this Letter.

Our results show that $\sim 28\%$ of the ECAs have rotation periods < 5.5 hours. Thus, by scaling the fraction of binary asteroids produced by our model ($\sim 57\%$) by the fraction of rapidly rotating asteroids in the ECA population ($\sim 28\%$), we can conclude that $\sim 15\%$ of all kilometre-sized Earth-crossers capable of producing binary asteroids should have satellites generated by Earth's tidal forces, and this steady-state ECA population describes the population that strikes Earth.

To determine the fraction of this population forming doublet craters on Earth, we modelled impact encounters between binary asteroids and the Earth, again using an adaptive fifth-order Runge-Kutta numerical integrator⁸ which accounts for Earth's tidal forces, the asteroids' encounter and impact velocities, and

the trajectory and orientation of the components at impact¹. We discovered that, contrary to earlier speculation and intuition, planetary tides just before impact decrease (on average) the distance between components at impact. Although the differential gravitational pull of the Earth during final approach increases the distance between the components, the increased separation is often in a direction radial to the planetary surface, such that the components tend to fall near or on top of one another. Consequently, $\sim 10\%$ of the asteroids striking the Earth produce doublet craters (compared with the 15% of Earth-crossers that are well separated binaries), matching the percentage of doublet craters observed on Earth (Fig. 3).

No recent methodical survey has yet been attempted to find asteroid satellites among the Earth-crossing asteroid populations, though doublet craters, anomalous stellar occultation lightcurves, unusual asteroid lightcurves, elongated asteroids, and extremely slow asteroid rotation rates all have suggested they may be quite common¹⁴. Now, in the aftermath of the interaction of comet Shoemaker-Levy 9 with Jupiter and the discovery¹⁵ of Dactyl orbiting 243 Ida (though Dactyl's small diameter and orbit around Ida, if not highly eccentric, suggests it would not create a doublet crater at impact), our results suggest that the time may be ripe for a new search. We recommend that any search for asteroid satellites places emphasis on kilometre-sized Earth-crossers with short rotation periods.

We note that 433 Eros, the target of the NEAR mission, has a short rotation period (5.27 hours) and an elongated shape, suggesting that it may have passed near enough to the Earth in the past to have produced a satellite. Although no such satellite has been reported, in spite of an intensive pre-encounter observing campaign, our model suggests at least a 50% probability of finding a small satellite in orbit around 433 Eros. $\square$

Received 7 December 1995; accepted 22 March 1996.

- Melosh, H. J. & Stansberry, J. S. *Icarus* **94**, 171–179 (1991).
- Cook, C., Melosh, H. J. & Bottke, W. F. *Lunar Planet. Sci.* **XXVI**, 275–276 (1995); *Icarus* (submitted).
- Passey, Q. & Melosh, H. J. *Icarus* **42**, 211–233 (1980).
- Sekiguchi, N. *Moon* **1**, 429–439 (1970).
- Oberbeck, V. R. & Aoyagi, A. *J. geophys. Res.* **77**, 2419–2432 (1972).
- Aggarwal, H. R. & Oberbeck, V. R. *Astrophys. J.* **191**, 577–588 (1974).
- Asphaug, E. & Benz, W. *Nature* **370**, 120–124 (1994).
- Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. in *Numerical Recipes* (Cambridge Univ. Press, 1986).
- Asphaug, E. & Benz, W. *Icarus* (in the press).
- Binney, J. & Tremaine, S. *Galactic Dynamics* (Princeton Univ. Press, 1987).
- Chauvineau, B., Farinella, P. & Harris, A. W. *Icarus* **115**, 36–46 (1995).

12. Bottke, W. F., Nolan, M. C., Greenberg, R. & Kolvoord, R. A. in *Hazard Due to Comets and Asteroids* (eds Gehrels, T. & Matthews, M. S.) (Univ. Arizona Press, Tucson, 1995).
13. Lagerkvist, C.-I. & Claesson, A. in *Small Bodies in the Solar System and their Interactions with the Planets* (eds Rickman, H. & Valtanen, M.) (in the press).
14. Weidenschilling, S. J., Paolicchi, P. & Zappala, V. in *Asteroids II* (eds Binzel, R. P., Gehrels, T. & Matthews, M. S.) 643–658 (Univ. Arizona Press, Tucson, 1989).
15. Belton, M. J. S. & Carlson, R. *IAU Circ. No.* 5948 (1994).
16. Schmidt, R. M. & Housen, K. R. *Int. J. Impact Engng* **5**, 543–560 (1987).
17. Melosh, H. J. *Impact Cratering: A Geologic Process* (Oxford Univ. Press, New York, 1989).

ACKNOWLEDGEMENTS. We thank E. Asphaug, B. Chauvineau, C. Cook, D. Durda, P. Farinella, R. Greenberg, A. Harris, M. Nolan, D. Richardson and J. Scotti for useful discussions and reviews of this manuscript. This work was partially supported by NASA's Planetary Geology and Geophysics Program and from a Texaco Prize Fellowship at Caltech (W.F.B.).

Observation of magnetically induced transverse diffusion of light

G. L. J. A. Rikken* & B. A. van Tiggelen†

* Grenoble High Magnetic Field Laboratory, Max Planck Institut für Festkörperforschung/CNRS, BP166, F-38042 Grenoble Cedex 9, France
† Laboratoire de Physique Numérique des Systèmes Complexes, CNRS/Université Joseph Fourier, Maison des Magistères, BP166, F-38042 Grenoble Cedex 9, France

PHOTONS and electrons, despite their very different nature, show many similarities in their behaviour. Several photonic counterparts of established electronic phenomena—such as photonic energy bands¹, weak localization^{2–4}, and the quantization of⁵ (and fluctuations in^{6–8}) optical transmission—have now been observed. These similarities originate in the wave-like character of both photons and electrons. But unlike photons, electrons are also charged, and thus experience the Lorentz force in a magnetic field. This force leads to the well known Hall effect, in which the application of a magnetic field to an electron-transporting medium generates a new current (or voltage) perpendicular to the direction of both the original current and the applied magnetic field. Despite the absence of photonic charge, one of us has predicted⁹ that the propagation of light through a disordered, scattering medium should be similarly affected by a magnetic field, although the origin of the effect is very different. Here we report the experimental confirmation of this phenomenon.

The magneto-transverse light current in disordered media is due to the well known magnetically induced change in optical properties of the medium⁹. Commonly, such a change is studied in homogeneous media by measuring the so-called Faraday rotation¹⁰, that is, the rotation of the polarization vector of a light wave propagating along the magnetic field lines. The rotation angle equals the product of a material parameter (the Verdet constant V), the amplitude of the magnetic field (B) and the traversed length d . A related phenomenon with the same VBd dependence is the magnetic rotation of speckles¹¹. The predicted⁹ magneto-transverse light current is proportional to VB , but only appears in a geometry different from the one above. Quantitative predictions for the magneto-transverse light current could only be obtained in the Rayleigh limit, that is, for scatterers of size smaller than the optical wavelength. But there is no reason why such a current should be restricted to small scatterers because its physical interpretation is a magnetically induced anisotropy of scattering characteristics, such as cross-section. This point of view reveals an analogy with heat transport in gases subjected to a magnetic field (the Beenakker–Senftleben effect^{12–15}). Here a magneto-transverse heat flow originates from the anisotropy of the scattering cross-section of the colliding molecules, unlike the Hall effect, which originates from the magnetic-field-induced deflection of the electron trajectories between the scatterers. In this respect the magneto-transverse light current predicted in ref. 9 has more in common with the Beenakker–Senftleben effect than with the Hall effect.

Diffuse transport of electrons is described by a conductivity tensor, and heat transport in gases by a thermal conductivity tensor. In an analogous fashion, multiple scattering of light in disordered media can be described by a diffusion tensor D . In media that are on average isotropic, diffusion must be isotropic and the associated diffusion tensor must be diagonal. The presence of an external magnetic field was found theoretically to introduce anisotropy in diffusion, that is, off-diagonal components in the diffusion tensor⁹. Fick's diffusion law relates the photon current density $\mathbf{J}$, defined per unit area, to the local photon density

gradient $\nabla\rho$ according to

$$\mathbf{J} = -D \cdot \nabla\rho \quad (1)$$

The magnetically induced off-diagonal components of the diffusion tensor D generate a magneto-transverse current density $\Delta\mathbf{J}_\perp$, that propagates perpendicular to both the magnetic field $\mathbf{B}$ and the photon gradient $\nabla\rho$, expressed mathematically as $\Delta\mathbf{J}_\perp \propto \ell_\perp \mathbf{B} \times \nabla\rho$. The gradient $\nabla\rho$ results from diffuse propagation of photons in the scattering medium, and is antiparallel to the incident beam. The parameter $\ell_\perp$ expresses the coupling of the photons inside the scattering medium to the magnetic field. It is linearly proportional to the magnetic field and the Verdet constant V , and has the dimension of length. The Verdet constant can have positive or negative sign, and therefore the same holds true for $\ell_\perp$. The sign translates directly into the direction of the magneto-transverse light current, just as the sign of the charge carriers determines the sign of the Hall effect.

To observe the magneto-transverse effect, we used the geometry shown in Fig. 1. The scattering medium is contained in a cylinder with its axis along the x -direction, with radius R and length L . The magnetic field $\mathbf{B}$ is directed along the z -axis. On solving the diffusion equation numerically for the cylindrical geometry¹⁶ with appropriate boundary conditions, we get expressions for the magnetically induced photon current difference $\Delta I_\perp$. This quantity describes the magneto-transverse current density $\Delta\mathbf{J}_\perp$ integrated over the area of the light collector. The relative importance of magneto-transverse light scattering over normal light scattering is obtained through normalizing by the radially scattered photon current $I_\perp$ collected by the same detector. Calculations for our sample dimensions ($L/R = 2.6$) show that in the opaque regime

$$\frac{\Delta I_\perp}{I_\perp} \approx 5 \frac{\ell_\perp}{R} \quad (2)$$

and is independent of the mean free path of the light ℓ . Opaque here means that $\ell \ll L$. The value for the magneto-transverse mean free path $\ell_\perp$ extracted from the experiments described below, will be compared to the prediction⁹ for a dilute sample of Rayleigh scatterers,

$$\ell_\perp = \frac{3n_m n_s}{(n_s^2 - n_m^2)(n_s^2 + 2n_m^2)} \frac{VB\lambda^2}{\pi^2} \quad (3)$$

where n_m and n_s are the refractive indices of matrix and scatterer,

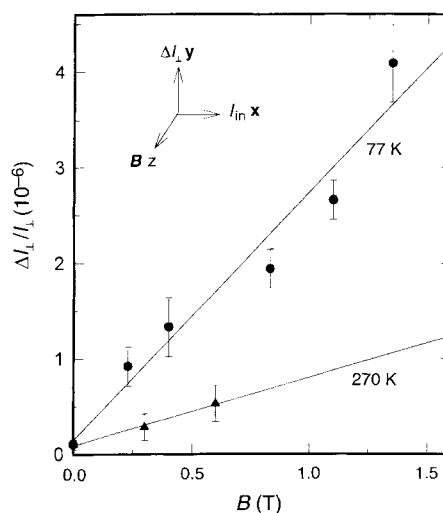


FIG. 1 Normalized magneto-transverse light current $\Delta I_\perp / I_\perp$ versus the amplitude B of the alternating magnetic field for 3.4 vol. % CeF_3 in glycerol, for two different temperatures. The straight line through the 77 K data points is a fit; the line through the 270 K data points was deduced from the 77 K measurements, using the known paramagnetic $1/T$ -dependence of the Verdet constant.

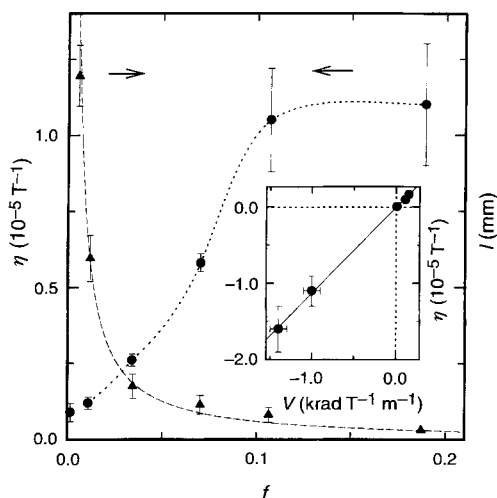


FIG. 2 Normalized magneto-transverse light current per unit magnetic field $\eta \equiv \Delta I_{\perp}/(I_{\perp}B)$ versus volume fraction f of CeF_3 in glycerol at $T = 77$ K. Also shown is the mean free path ℓ which was estimated from transmission measurements using an integrating sphere. Its error bars do not include a 30% systematic uncertainty in interpreting such measurements. The dashed line is a least-squares fit to an expected $1/f$ dependence, the dotted line is only meant to guide the eye. Inset, η versus the Verdet constant V of different scatterers in the opaque regime (from left to right; EuF_2 , CeF_3 , Al_2O_3 , TiO_2 and ZnS , all suspended in glycerol, having a Verdet constant of $V = 6.3 \text{ rad T}^{-1} \text{ m}^{-1}$). The sign of the magneto-transverse current was taken into account by means of the lock-in phase angle ϕ . Straight line is a fit to the data points.

and λ the wavelength *in vacuo*. We note that the magneto-transverse length $\ell_{\perp}$ and hence the normalized difference $\Delta I_{\perp}/I_{\perp}$ do not depend on the volume fraction f of the scattering particles, once the system is in the optically thick (opaque) regime. This feature, as well as the direction of the magneto-transverse flux are two stringent experimental tests. For $V < 0$, this direction is predicted to be along the positive y -axis in Fig. 1 inset. On the basis of our solution of the diffusion equation, we expect that on going from the optically thick towards the optically thin regime, the numerical factor in equation (2) gradually decreases. For optically thin media ($\ell > L$) the diffusion approach is not valid, and no clear statements can be made.

Light-scattering experiments were performed on mixtures of glycerol ($n_m = 1.48$) and CeF_3 powder ($n_s = 1.62$). The particle size of the powder is roughly $a \approx 2 \mu\text{m}$. These materials were selected because glycerol has a small Verdet constant whereas CeF_3 has a large Verdet constant which in addition can be varied by changing the temperature¹⁷. As a result the magnetic field predominantly affects the optical properties of the CeF_3 scatterers. The mixtures were contained in a cylindrical cavity in Perspex ($R = 0.5 \text{ mm}$, $L = 1.3 \text{ mm}$). One end of the cylinder was homogeneously illuminated by light from an argon-ion laser through an optical fibre. The light incident on the end of the cylinder was found to be unpolarized. The wavelength of the light is 457 nm , that is, the particles are considerably larger than the wavelength. Quantitative agreement between the theory cited above and experiment cannot therefore be expected. Scattered light was collected by two opposing fibres perpendicular to both the cylinder axis and the magnetic field and was detected by photodiodes. The magnetic field was supplied by an electromagnet, generating a 40-Hz alternating field with a maximum amplitude of 1.35 T . The difference between the photodiode signals was detected by a lock-in amplifier. In this way, the lock-in output is proportional to the magneto-transverse current $\Delta I_{\perp}$. This output was normalized by the transversely scattered light current $I_{\perp}$ as in equation (2).

The results shown in Fig. 1 confirm a linear field dependence of the normalized magneto-transverse current $\Delta I_{\perp}/I_{\perp}$. The temperature (T) dependence of the two slopes is consistent with the known paramagnetic $1/T$ -dependence of the Verdet constant¹⁷. The other optical parameters hardly change with temperature; Fig. 1 therefore proves that the effect is due to the magnetically induced changes in the optical properties of the scatterers, linear in V and B . It is therefore convenient to consider from now on the slope $\eta \equiv \Delta I_{\perp}/(I_{\perp}B)$ as a measure of the magneto-transverse current. From the lock-in phase angle between the oscillating signal $\Delta I_{\perp}(t)$, where t is time, and the oscillating magnetic field $B(t)$ we conclude the observed sign of the magneto-transverse current to be consistent with the one predicted above. The parameter η as well as the mean free path ℓ were determined as a function of the volume fraction f of scatterers (Fig. 2). For $f \geq 10\%$, the data given in Fig. 2 show that $\ell \ll L$ and that the normalized transverse current is independent of the volume fraction of the scatterers, as predicted by equation (2) for the opaque regime. The experimental value for η corresponds, by equation (2), to $\ell_{\perp}/B \approx 1 \text{ nm/T}$. The prediction, equation (3), for Rayleigh scatterers is $\ell_{\perp}/B \approx 0.1 \text{ nm/T}$. We attribute the difference to the fact that our mixtures do not contain Rayleigh scatterers. In the dilute regime $f \leq 1\%$ we find that the mean free path is larger than the cylinder. Extrapolating our experimental data towards vanishing volume fraction, that is, vanishing scattering, we obtain a non-zero value for η . This remarkable phenomenon is now under study. We also determined the slope η in the optically thick regime for four other materials (Fig. 2 inset). The observed linear relationship $\eta(V)$ is somewhat surprising because equation (3) for $\ell_{\perp}$ predicts a strong dependence on the index of refraction n_s , ranging from 1.6 to 2.7. Again, this discrepancy may be attributed to the non-Rayleigh character of our scatterers. We emphasize that both signs of the magneto-transverse current have been observed, and found to be consistent with the sign of the Verdet constant of the scatterers as predicted by theory.

These experiments have firmly established the magneto-transverse light current, and confirmed its qualitative features (as predicted in ref. 9) for particles smaller than the wavelength of the light. The scattering particles in our experiments are larger than the wavelength. From our experiments we conclude that the magneto-transverse light current for mixtures containing large particles is much bigger than predicted in ref. 9, and shows little dependence on refractive index. Evidently, a theory for large particles is needed to make a quantitative comparison. We have established empirically that the magneto-transverse current may be used to determine the sign and the magnitude of the Verdet constant of powdered materials. □

Received 22 September 1995; accepted 8 March 1996.

1. Yablonoitch, E. in *Analogies in Optics and Micro Electronics* Ch. 8 (eds Van Haeringen, W. & Lenstra, D.) 117–134 (Kluwer, Dordrecht, 1990).
2. Van Albada, M. P. & Lagendijk, A. *Phys. Rev. Lett.* **55**, 2692–2695 (1985).
3. Wolf, P. E. & Maret, G. *Phys. Rev. Lett.* **55**, 2696–2699 (1985).
4. Erbacher, F., Lenke, R. & Maret, G. *Europhys. Lett.* **21**, 551–556 (1993).
5. Montie, E. A., Cosman, E. C., 't Hooft, G. W., Van der Mark, M. B. & Beenakker, C. W. J. *Nature* **350**, 594–595 (1991).
6. Garcia, G. & Genack, A. Z. *Phys. Rev. Lett.* **63**, 1678–1681 (1989).
7. Van Albada, M. P., De Boer, J. F. & Lagendijk, A. *Phys. Rev. Lett.* **64**, 2787–2790 (1990).
8. Berkovits, R. & Feng, S. *Phys. Rep.* **238**, 135–172 (1994).
9. Van Tiggelen, B. A. *Phys. Rev. Lett.* **75**, 422–424 (1995).
10. Landau, L. D., Lifshitz, E. M. & Pitaevskii, L. P. *Electrodynamics of Continuous Media* Sect. 97 (Pergamon, Oxford, 1984).
11. Darshit, M. Ya., Zhirgalova, I. V., Zel'dovich, B. Ya. & Kundikova, N. D. *JETP Lett.* **59**, 763–765 (1994).
12. Kagan, Yu. & Maksimov, L. *JETP* **14**, 604–610 (1962).
13. Korving, J., Hulsman, H., Knaap, H. F. P. & Beenakker, J. J. M. *Phys. Lett.* **21**, 5–6 (1966).
14. Hermans, L. J. F., Fortuin, P. H., Knaap, H. F. P. & Beenakker, J. J. M. *Phys. Lett.* **25A**, 81–82 (1967).
15. Mazur, E., Hijken, H. J. M., Hermans, L. J. F. & Beenakker, J. J. M. *Physica A* **123**, 412–427 (1984).
16. Crank, J. *The Mathematics of Diffusion* (Clarendon, Oxford, 1990).
17. Leycuras, C., LeGall, H., Guillot, M. & Marchand, A. J. *appl. Phys.* **55**, 2161–2163 (1984).

ACKNOWLEDGEMENTS. We thank R. Maynard, P. Wyder, A. Lagendijk, G. Maret and R. Lenke for helpful discussions.

Control of crystal phase switching and orientation by soluble mollusc-shell proteins

A. M. Belcher^{*†}, X. H. Wu[‡], R. J. Christensen[‡],
P. K. Hansma[§], G. D. Stucky^{*‡} & D. E. Morse^{†||}

Materials Research Laboratory and * Department of Chemistry,
† Marine Biotechnology Center, ‡ Department of Materials,
§ Department of Physics, || Department of Molecular,
Cellular and Developmental Biology, University of California,
Santa Barbara, California 93106, USA

In the initial stages of the biomineralization of abalone shells, a primer layer of oriented calcite crystals grows on a nucleating protein sheet^{1,2}. The deposition of this primer is followed by an abrupt transition to *c*-axis-oriented crystals of aragonite, another crystalline form of calcium carbonate. The formation of each of the two crystal types is accompanied by the synthesis of specific polyanionic proteins¹⁻³, suggesting that cooperative interactions between these proteins and the inorganic ions during crystal nucleation and growth control the phase of the deposited mineral and that differential expression of the proteins allows the organism to induce phase changes. It is known that soluble shell proteins can control crystal morphology⁴⁻¹⁰, but it has been suspected that the switch in phase—from calcite to aragonite—might require the deposition of a new nucleating protein sheet. Here we describe *in vitro* studies of the crystallization of calcium carbonate in the presence of soluble polyanionic proteins extracted from abalone shell. We find that these proteins alone are sufficient to control the crystal phase, allowing us to switch abruptly and sequentially between aragonite and calcite without the need for deposition of an intervening protein sheet. These results show that soluble organic components can exert greater control over hierarchical biomineral growth than hitherto suspected, offering the prospect of similar phase control in materials chemistry.

The molluscan shell is a microlaminate composite of mineral and biopolymers exhibiting exceptional nanoscale regularity³, and a strength ~ 3,000 times greater than that of the crystals themselves¹¹. Although the integral proteins typically comprise < 2 wt% of the shell, they determine the structural organization and properties of the mineralized composite. We previously found that a 'flat pearl' closely duplicating the composition and organization of the native shell can be biofabricated on glass or mica coverslips inserted between the mantle (secretory epithelium) and shell of *Haliotis rufescens* (red abalone; gastropod mollusc)¹. Biomineralization of the flat pearl is spatially and temporally regulated in a developmental sequence closely parallel to that at the growth front of the natural shell^{1,2}. A nucleating sheet of protein is first secreted over the substrate; a layer of calcite (CaCO₃) crystals then is nucleated in specific (104) orientation on this protein sheet, and after the nucleation centres have grown to confluence, there is an abrupt transition to the production of *c*-axis-oriented aragonite (a second polymorph of CaCO₃) organized in multiple layers of interdigitating flat polygonal crystals (~ 0.5 µm thick × ~ 5–10 µm) to form nacre (mother-of-pearl)^{1,2}. Transverse sections of shell reveal similar, highly oriented calcite spherulites that grow from apparently identical nucleating protein sheets deposited interstitially in the shell^{2,12,13}.

Previous investigations demonstrated that proteins from shell and other biomineralized composites can control the morphology⁴⁻¹⁰ of CaCO₃ crystals. Falini *et al.* recently showed that soluble shell proteins can determine the polymorph of crystals grown on a heterologous substrate of squid chitin and silkworm fibroin¹⁴, in contrast to earlier work that appeared to demonstrate polymorph control by networks of insoluble proteins extracted from shells¹⁵.

The calcitic and aragonitic composites contain distinct, mineral-phase-specific populations of crystal-associated, polyanionic proteins that can be solubilized by demineralization (Fig. 1A). Gel electrophoresis under denaturing conditions demonstrates that the aragonitic composite contains three major species detectable with a cationic carbocyanin stain^{1,16} at relative molecular mass (*M_r*) 16K, 20K and 31K; in contrast, the calcitic composite contains six major proteins with *M_r* of 14K, 28K, 35K, 48K and 55K (data not shown). The 16K aragonitic species is intracrystalline (Fig. 1A). Amino-acid compositions, electrophoretic mobilities under non-denaturing conditions, and staining with the cationic carbocyanin dye all confirm that these proteins are polyanionic^{3,16}.

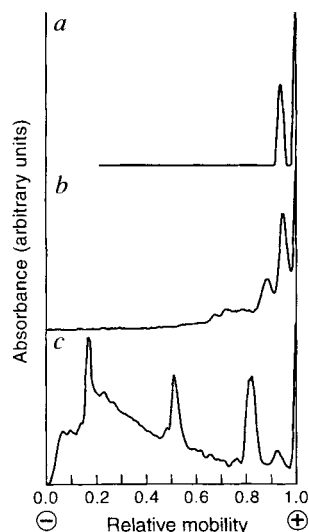
The nucleating protein sheet and soluble mineral-specific polyanionic proteins control CaCO₃ crystal growth *in vitro* (Fig. 1B). Crystal nucleation density was always higher, and the crystals more organized and uniformly spaced, on the protein sheet than on glass. Crystals grown in the absence of soluble protein (Fig. 1B, a) exhibit the characteristic rhombohedral morphology of calcite. These crystals were identical to those grown in the presence of bovine serum albumin (added in solution) under otherwise identical conditions. In contrast, crystals grown from a solution containing the polyanionic proteins extracted from the calcitic composite of the shell exhibit the characteristic spherulitic morphology of calcite (Fig. 1B, b).

Crystals grown in the presence of the polyanionic proteins solubilized from the aragonitic composite displayed a very different morphology, forming needles of aragonite in the plane of the nucleation layer (Fig. 1B, c). Crystal phase, morphology and orientation were verified by Raman microprobe and optical crystallography; electron diffraction revealed the aragonite orthorhombic lattice and diffraction extinctions consistent with the *Pmcn* space group.

Crystals grown from a mixture of proteins from the aragonitic and calcitic composites proved to be aragonite. The predominant crystal type was flat polycrystalline plates that exhibit uniform extinction under cross-polarized light (Fig. 1B, d). The individual plates were shown by electron diffraction to be aragonite with a zone axis of $\langle 001 \rangle$ as found in the nacre of the shell and flat pearl. Transmission electron microscopy and electron diffraction further revealed that these truncated *c*-axis plates grew in crystallographically oriented stacks resembling those in biogenic nacre (Fig. 2), despite the fact that the insoluble protein matrix surrounding the tablets of biogenic nacre of *H. rufescens* had not been added.

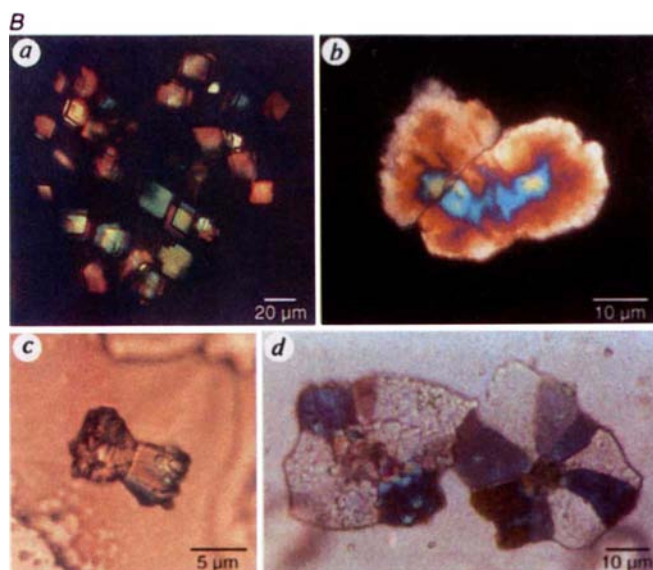
Addition of mineral-specific proteins proved sufficient to induce the abrupt calcite-to-aragonite transition characteristic of abalone shell and pearl biomineralization^{1,2}. Rhombohedral calcite crystals first were grown on the nucleating protein sheet from a saturated solution of calcium carbonate in the absence of soluble proteins. This solution then was removed and replaced with a fresh solution containing the same inorganics together with the soluble polyanionic proteins extracted either from the calcitic or aragonitic composites of the shell. Optical polarized microscopy and scanning electron microscopy revealed that the resulting new crystal growth was specific to the faces of the calcite rhombohedron and that the two families of proteins produced different crystal morphologies (Fig. 3a–c). X-ray diffraction studies (Fig. 4, showing calcite (104) and (113) diffraction peaks) and Raman microprobe analyses of these crystals reveal that the inorganically grown rhombohedrons are calcite, whereas the needles grown on this calcite (Fig. 3a, b, d) are aragonite (compare Fig. 4, showing aragonite (111) and (130) diffraction peaks). These data show that the abrupt transition from (104) calcite to aragonite seen in the biofabrication of the shell and flat pearl can be duplicated simply by addition of the soluble aragonitic proteins to the crystal growth solution. Additional evidence for the stereospecificity of protein-directed mineralization is seen in the case of extensive crystal overgrowth on the rhombohedron faces in the presence of these proteins (Fig. 3e, f), generating

FIG. 1 Effect of shell proteins on crystal growth. A, Non-denaturing gel electrophoretic separation of polyanionic proteins from shell: a, intracrystalline aragonitic protein; b, total proteins isolated from aragonitic portion of the shell (two major species and two minor species); c, total proteins isolated from calcitic portion of the shell (nine species seen). B, Crystals grown on the nucleating protein sheet in the presence or absence of polyanionic proteins isolated from the calcitic or aragonitic domains of shell (optical microscopy under crossed polarized light). a, Calcite rhombohedrons grown in the absence of polyanionic proteins. Calcite was identified both by X-ray diffraction revealing the calcite (104) reflection and by Raman microprobe analysis ($\nu_1 = 1,086.0 \text{ cm}^{-1}$; $\nu_4 = 712.3 \text{ cm}^{-1}$; values for geological calcite (104): $\nu_1 = 1,086.2 \text{ cm}^{-1}$;



$\nu_4 = 712.0 \text{ cm}^{-1}$; biological calcite (001): $\nu_1 = 1,085.4 \text{ cm}^{-1}$; $\nu_4 = 712.4 \text{ cm}^{-1}$). b, Calcite spherulitic crystals grown in presence of calcitic polyanionic proteins; confirmation by X-ray diffraction revealing the (104) reflection and by Raman microprobe analysis ($\nu_1 = 1,086.2 \text{ cm}^{-1}$; $\nu_4 = 712.0 \text{ cm}^{-1}$). c, Aragonite needles ($\sim 10 \mu\text{m}$) grown in presence of aragonitic polyanionic proteins. Aragonite was verified by Raman microprobe, revealing a red shift in the ν_1 vibrational mode ($\nu_1 = 1,083.5 \text{ cm}^{-1}$; values for biological aragonite (001): $\nu_1 = 1,084.9 \text{ cm}^{-1}$; $\nu_4 = 701.4$ and 705.6 cm^{-1} ; geological aragonite (001): $\nu_1 = 1,085.3 \text{ cm}^{-1}$; $\nu_4 = 701.2$ and 705.6 cm^{-1}); crystal phase and orientation were verified by electron diffraction, demonstrating that the red shift in the Raman data is correlated with the orientation discussed in the text. d, Aragonite grown in presence of both calcitic and aragonitic polyanionic proteins.

METHODS. Calcitic and aragonitic portions of the shell were mechanically separated and finely ground; X-ray diffraction was used to verify purity of mineral phase before protein isolation. The composites were demineralized with either acetic acid or EDTA, and the solubilized proteins dialysed exhaustively against water; comparable results were obtained with either



method. Treatment of the aragonitic nacre by sonication in the presence of 4–6% hypochlorite to disaggregate the nacre tablets (as verified by electron microscopy) and to oxidatively destroy the extracrystalline proteins before demineralization revealed that one of the aragonitic proteins was protected from the hypochlorite, and thus is intracrystalline. Electrophoresis under non-denaturing conditions was in 10–20% gradient polyacrylamide gel in Tris-tricine buffer; proteins were stained with 1-ethyl-2-[3-(1-ethyl-naphthol [1,2-d]-thiazolin-2-ylidene)-2-methylpropenyl]-naphtho[1,2-d]thiazolium bromide and scanned with a PDI optical densitometer. Nucleating protein sheets were purified from acid-demineralized shell; sheets from flat pearl and shell are virtually identical in amino acid composition, and particularly rich in proline, tyrosine and glycine (19.5, 16.0 and 16.3 residue-%, respectively); autofluorescence and absorption spectra, and resistance of the two-dimensional sheets to a wide range of denaturing agents and proteolytic enzymes, indicate a high degree of covalent cross-linking. Crystals were grown from CaCl_2 (12 mM) and NH_4HCO_3 (12 mM) at pH 8.0 on the purified nucleating protein sheets; soluble polyanionic proteins or bovine serum albumin (control) were added to a final concentration of $2 \mu\text{g ml}^{-1}$.

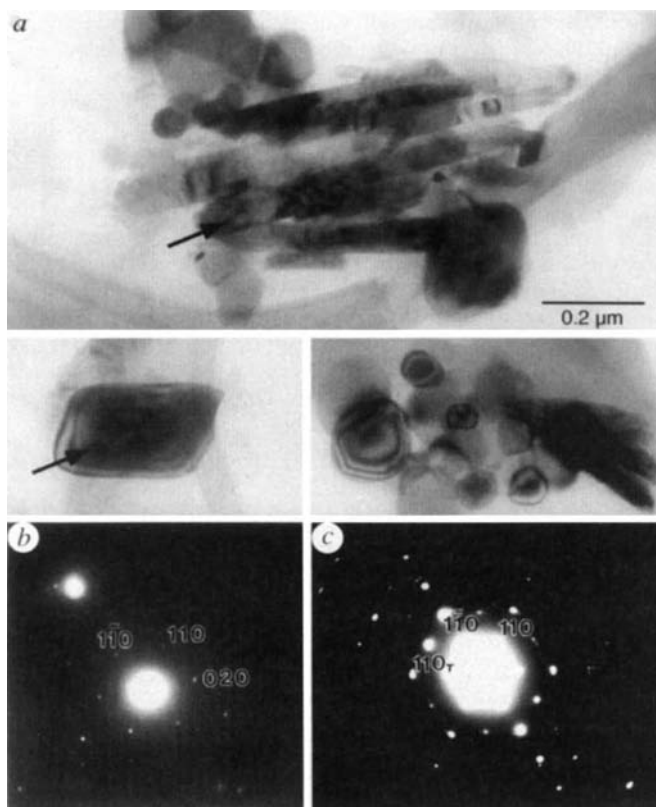


FIG. 2 Crystals grown in the presence of the nucleating protein sheet and aragonitic or calcitic polyanionic proteins. a, Transmission electron microscopy reveals aragonite with stacked plate morphology similar to the 'stack of coins' morphology of aragonitic nacre and the same crystallographic orientation. Electron microscopy performed with a Jeol 2000 fx instrument at 200 kV. In the top portion, crystal plates are viewed perpendicular to the c-axis; in the lower portions, others are viewed parallel to the c-axis. b, c, Diffraction patterns from selected areas indicated by arrows in a: b, diffraction (from region indicated by arrow in lower left of a) showing plate is aragonite with (001) zone axis; c, diffraction (from region indicated by arrow in top portion of a) showing twinning pattern of plate in another c-axis-oriented stack.

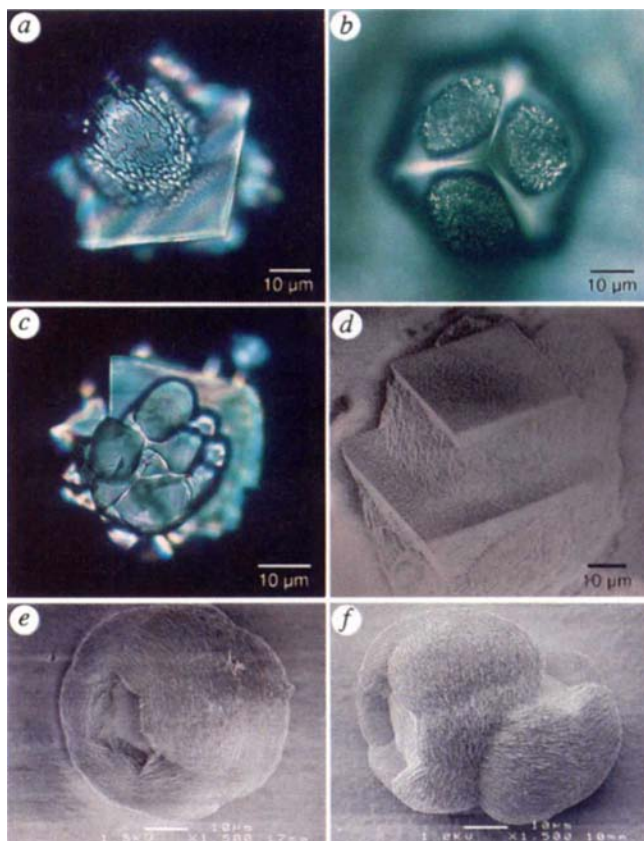
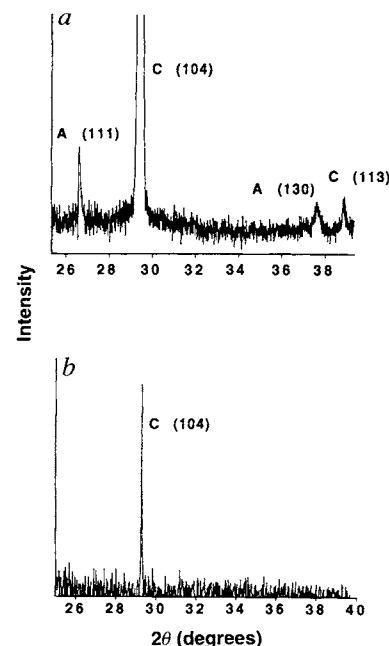


FIG. 3 Inorganic growth of calcite rhombohedrons with subsequent overgrowth in the presence of polyanionic proteins. Calcite rhombohedrons were grown on the nucleating protein sheet from solutions of 12 mM CaCl_2 and NH_4HCO_3 . After rhombohedrons were formed, the growth solution was removed and replenished with a fresh solution with or without the addition of polyanionic proteins (final concentration, $2 \mu\text{g ml}^{-1}$). The resulting crystals were viewed by optical microscopy using crossed polarizers (a–c) or scanning electron microscopy (d–f): a, b, Addition of aragonitic polyanionic proteins causes aragonite overgrowth as needles on the rhombohedral faces; in b, the view is down the c-axis of the calcite rhombohedron. c, Addition of calcitic polyanionic proteins causes calcite overgrowth on rhombohedral faces. d, Scanning electron microscopy of sequential phase switching induced by the sequential addition and removal of aragonitic polyanionic proteins. e, f, Extensive overgrowth of calcite rhombohedrons in the presence of aragonitic proteins generates composite crystals with sixfold symmetry about specific faces of the rhombohedrons; no overgrowth is seen on the rhombohedral c axis. Electron microscopy was with a Jeol JSM 6300F equipped with a cold-cathode field-emission source; the sample was uncoated and examined with a beam energy of 1.0 kV.

structures with sixfold symmetry with no overgrowth observed in the [001] (c-axis) direction. Sequential transition of calcite to aragonite and back once again to calcite (Fig. 3d) occurs when the soluble aragonitic proteins are depleted, permitting formation of

FIG. 4 X-ray diffraction demonstrates calcite to aragonite transition. Reflections specific to aragonite (peaks labelled 'A') and calcite ('C') are identified. a, 2θ scan identifies new aragonite needles ((111) $2\theta = 26.4^\circ$; (130) $2\theta = 38.1^\circ$) overgrowing the calcite rhombohedrons ((104) $2\theta = 29.3^\circ$; (113) $2\theta = 39.5^\circ$); crystals shown in Fig. 3a, b. b, 2θ scan of calcite rhombohedrons grown from saturated solution in the absence of soluble proteins; the (104) reflection of calcite is evident; crystals as shown in Fig. 1B. a, X-ray diffraction studies were carried out on a Philips 4-circle diffractometer.



calcite again as the stable phase. Subsequent addition of aragonitic proteins then results in the growth of new aragonite needles on the second layer of calcite rhombohedrons. The presence or absence of the aragonite-specific polyanionic proteins is thus sufficient to switch the polymorph of growing crystals at the micrometre scale, without the addition of magnesium or a change in temperature or pressure.

The results reported here demonstrate that only the polyanionic proteins from abalone shell and flat pearl are needed to provide stereospecific control of nucleation, crystal orientation, polymorph phase, and switching back and forth between phases, while the nucleating protein sheet determines the growth orientation of the calcite primer, all with high fidelity to the biological system. These results can thus account for the transition between calcite and aragonite observed in shell and flat pearl synthesis^{1–3}. Surprisingly, addition of soluble shell proteins is sufficient to promote formation of the stacked thin plates of crystallographically oriented aragonite crystals resembling those in abalone nacre. We suggest that the polyanionic proteins contribute to the higher-order crystal plate morphology, while the precise thickness of the nacre plates formed during biomineralization may be governed by the fixed interlamellar spacing¹⁷ of the insoluble proteins of the matrix. These results show that regulation of nucleation, growth, morphology and aggregation (phase) does not always require condensed, pre-organized organic arrays, in contrast to a commonly accepted central tenet of biomineralization that these parameters are controlled by organized assemblies of organic macromolecules^{9,18,19}. □

Received 25 October 1995; accepted 21 March 1996.

1. Fritz, M. et al. *Nature* **371**, 49–51 (1994).
2. Zaremba, C. M. et al. *Chem. Mater.* **8**, 679–690 (1996).
3. Morse, D. E. et al. *Mater. Res. Soc. Symp. Proc.* **292**, 59–67 (1993).
4. Addadi, L. & Weiner, S. in *Biomineralization: Chemical and Biochemical Perspectives* (eds Mann, S., Webb, J. & Williams, R. J. P.) (VCH, New York, 1989).
5. Aizenberg, J. et al. *J. Cryst. Growth* **142**, 156–164 (1994).
6. Aizenberg, J. et al. *FASEB J.* **9**, 262–268 (1995).
7. Albeck, S. et al. *J. Am. chem. Soc.* **115**, 11691–11697 (1993).
8. Heywood, B. R. & Mann, S. *Adv. Mater.* **6**, 9–20 (1994).
9. Mann, S. et al. *Science* **261**, 1286–1292 (1993).
10. Greenfield, E. M. & Chrenshaw, M. A. in *Origin, Evolution, and Modern Aspects of Biomineralization in Plants and Animals* (ed. Crick, R. E.) 303–308 (Plenum, New York, 1989).
11. Currey, J. D. *Proc. R. Soc. Lond. B* **196**, 443–463 (1977).
12. Erasmus, J., Cook, P. A. & Sweij, N. J. *Shellfish Res.* **13**, 493–501 (1994).

13. Shepherd, S. *Aust. J. mar. Freshwat. Res.* **46**, 607–615 (1995).
14. Falini, G. et al. *Science* **271**, 67–69 (1996).
15. Wilbur, K. M. & Watabe, N. *Ann. N.Y. Acad. Sci.* **109**, 82–112 (1963).
16. Cirioli, M. A. & Morse, D. E. *J. Comp. physiol.* **B157**, 717–729 (1988).
17. Nakahara, H., Bevilander, G. & Kakei, M. *Venus* **41**, 33–46 (1982).
18. Weiner, S. *Crit. Rev. Biochem.* **20**, 365–408 (1986).
19. Mann, S. *Nature* **132**, 119–124 (1988).

ACKNOWLEDGEMENTS. We thank W. Wise, Y. Li, J. S. Speck, M. Cornish, R. Day and R. Yuen for their generous assistance and helpful suggestions, and L. J. Friesen for the optical microphotography. This work was supported by the MRL Program of the US NSF, the US Office of Naval Research (D.E.M., G.D.S. and P.K.H.) and the Materials Research Division of the US NSF (G.D.S., D.E.M. and P.K.H.). A.M.B. gratefully acknowledges support from a Parsons Fellowship awarded through the Materials Research Laboratory, University of California at Santa Barbara.

CORRESPONDENCE should be addressed to D.E.M.

Synthesis of long prebiotic oligomers on mineral surfaces

James P. Ferris*, Aubrey R. Hill Jr†, Rihe Liu†
& Leslie E. Orgel†

* Department of Chemistry, Rensselaer Polytechnic Institute, Troy, New York 12180, USA

† The Salk Institute for Biological Studies, PO Box 85800, San Diego, California 92186, USA

Most theories of the origin of biological organization assume that polymers with lengths in the range of 30–60 monomers are needed to make a genetic system viable¹. But it has not proved possible to synthesize plausibly prebiotic polymers this long by condensation in aqueous solution, because hydrolysis competes with polymerization. The potential of mineral surfaces to facilitate prebiotic polymerization was pointed out long ago². Here we describe a system that models prebiotic polymerization by the oligomerization of activated monomers—both nucleotides and amino acids. We find that whereas the reactions in solution produce only short oligomers (the longest typically being a 10-mer), the presence of mineral surfaces (montmorillonite for nucleotides, illite and hydroxylapatite for amino acids) induces the formation of oligomers up to 55 monomers long. These are formed by successive ‘feedings’ with the monomers; polymerization takes place on the mineral surfaces in a manner akin to solid-phase synthesis of biopolymers^{3,4}.

Solid-phase synthesis depends on the ease with which growing chains are separated from spent reagents in solution by washing. The attachment to the resin must be irreversible, or nearly so. If the strength of adsorption of a polymer to a surface increases with length, sufficiently-long polymers will be adsorbed almost irreversibly and can then be extended indefinitely by the prebiotic equivalent of solid-phase synthesis. Our experimental approach is simple. A mineral is incubated with an activated monomer until a family of short oligomers are formed. The solid is separated by centrifugation and the activated monomers are replaced. This process is repeated as often as necessary for the production of oligomers of length 40 or more. Finally, the solid is eluted with a solution of sodium pyrophosphate and the eluate is analysed.

Montmorillonite is a very effective catalyst for the oligomerization of nucleoside 5'-phosphorimidazolides^{5,6}. We therefore studied the elongation of the decanucleotide [³²P]dA(pdA)₈pA adsorbed on Na⁺-montmorillonite by daily additions of ImpA (Fig. 1). Polyadenylates containing more than 20 monomers were formed after ‘feeding’ twice with ImpA, with the main products being 11–14-mers (Fig. 2). Polynucleotides containing more than 50 monomers were formed after 14 feedings, with the principal oligomeric products containing 20–40 monomer units (Fig. 2). A small amount of the pyrophosphate-capped primer was the only product observed after a 12-day reaction in the absence of montmorillonite (data not shown).

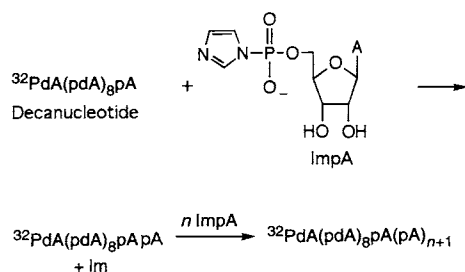


FIG. 1 Elongation of a decanucleotide by reaction with ImpA: Im is imidazole, pA is adenosine-5'-phosphate, pdA is 3'-deoxyadenosine-5'-phosphate, and ³²P indicates a radioactively labelled phosphate group.

The oligomerization of glutamic acid in the presence of the condensing agent carbonyl diimidazole is a very efficient reaction that proceeds via an *N*-carboxy anhydride intermediate⁷. The products of a reaction using a relatively dilute solution of the amino acid are illustrated in Fig. 3a. Oligomers up to the 10-mer can be detected, but the great majority of the products are shorter than the 5-mer. If the reaction is carried out in the presence of illite, the shorter oligomers remain in the supernatant and longer oligomers adsorb to the mineral. There is no evidence for catalysis by the mineral. The products eluted from the mineral after 10, 25 and 50 feedings are illustrated in Fig. 3b, c and d, respectively. After 50 feedings oligomers up to at least the 55-mer can be detected, and the bulk of the adsorbed product is in the size range 30–50.

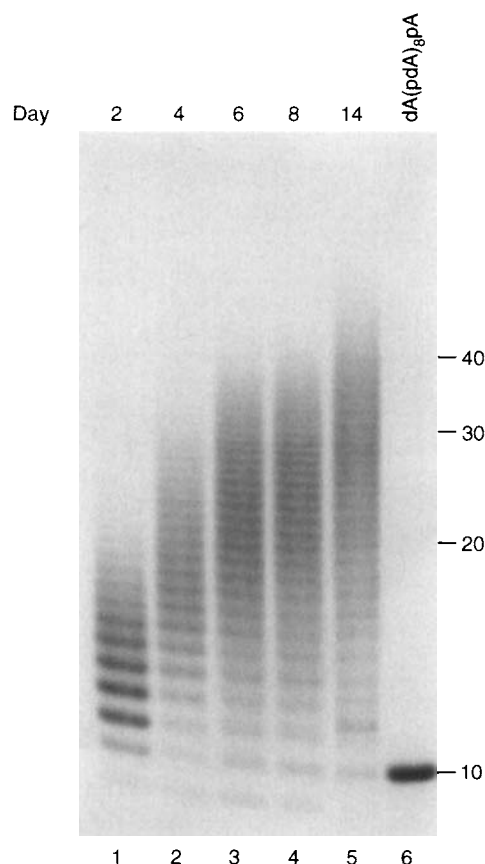
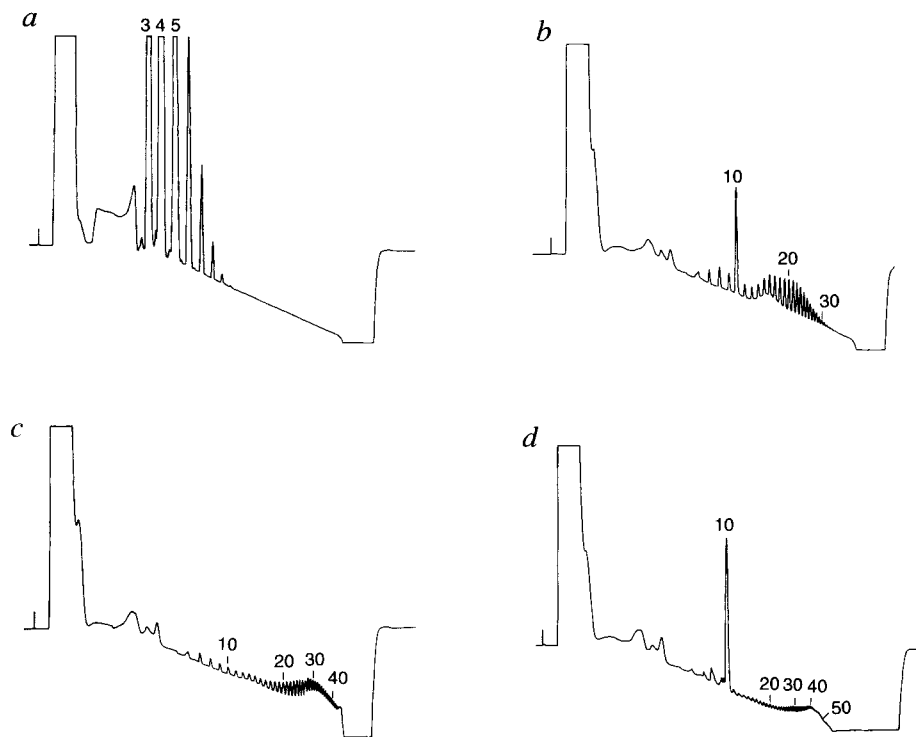


FIG. 2 Gel electrophoresis of the polyadenylates formed by the elongation of [³²P]dA(pdA)₈pA. Migration positions of size markers (for 10-, 20-, 30- and 40-mers) are shown on the right. dA(pdA)₈pA is prepared on an Applied Biosystems 391 DNA synthesizer initiated by adenosine attached to porous glass beads (Milligen). The deblocked oligomer is phosphorylated with a mixture of T4-polynucleotide kinase (New England Biolabs) and [³²P]ATP (Amersham) and the product is deionized and purified on a Nensorb (duPont) column. The primer migrated as one band on gel electrophoresis in a 20% polyacrylamide denaturing gel using a buffer of 0.09 M Tris-borate and 0.002 M EDTA (TBE) buffer, pH 8. The primer (~500,000 c.p.m.) is dissolved in 40 µl buffer (0.1 M HEPES, 0.1 M NaCl, 0.075 M MgCl₂, pH 8) and added to 2 mg Na⁺-montmorillonite in a Gelman 4937 Z-spin filtration unit containing a 0.45-µm filter. To this is added sufficient ImpA to give a 15 mM solution. The reaction mixture is vortexed and allowed to stand for one day at 25 °C. The tube is then centrifuged to remove the aqueous phase from the montmorillonite and then washed with 40 µl of the buffer reaction solution. Fresh ImpA and buffer salt (total volume 40 µl) are added to the clay-oligomer complex in the filtration tube and the reaction allowed to proceed for another day. Six reaction tubes are prepared and, after appropriate numbers of incubations, gel products are eluted from the montmorillonite by 3 × 40 µl washes of 0.1 M pyrophosphate (pH 9). The combined washes are deionized on a Nensorb column and analysed by gel electrophoresis on a 20% gel using TBE buffer.

FIG. 3 Polymerization of glutamic acid (Glu) on illite. a, Control reaction in the absence of illite; b, eluate after 10 feedings; c, after 25 feedings; d, after 50 feedings. A marker of authentic (Glu)₁₀ is added to the samples in b and d. Illite (10 mg or 20 mg) is weighed in a microfuge tube. As illite cannot be divided into aliquots accurately, a separate tube is prepared for each reaction sequence. A solution of activated monomer is prepared by adding 20 mM L-Glu (pH 8.0) to a 2.5-fold excess of dry, solid carbonyl diimidazole at 0 °C. The solution is kept at 0 °C for 1 min, after which 50 µl aliquots are pipetted into the illite-containing reaction tubes. The tubes are vortexed to suspend the illite and then tumbled at room temperature. After 6 h the tubes are spun for 1 min, the supernatant is removed, and an aliquot of freshly prepared activated glutamic acid is added. The tubes are vortexed to resuspend the illite and tumbled for 18 h. The tubes are then spun again for 1 min, the supernatant removed, and activated glutamic acid is again added. The sequence is repeated, with alternating incubations of 6 and 18 h until the illite has been 'fed' an appropriate number of times. The reaction tubes are then spun for 1 min, and the supernatant set aside. The solid is washed four times with 250 µl water. It is then eluted with each of three 500-µl aliquots of 20 mM K₄P₂O₇ for 15 min. The eluates are combined and analysed by HPLC on RPC-5 using a perchlorate gradient (pH 8; 0–0.03 M in 40 min).



We have also used a water-soluble carbodiimide, *N*-(3-dimethylaminopropyl)-*N*-ethylcarbodiimide hydrochloride (EDAC) as a condensing agent for the polymerization of aspartic acid on the surface of hydroxylapatite. In this system both of the carboxyl groups of the amino acid are activated, so branching is possible. Because this leads to a complex family of isomers for each molecular weight, individual peaks corresponding to longer oligomers are poorly resolved in the high performance liquid chromatography (HPLC) traces. Figure 4 shows elution profiles of the products formed in homogeneous solution, and the compositions of the material eluted from the mineral after different numbers of 'feedings'. Clearly, very large amounts of material of high molecular weight are accumulated. After 30 feedings, most of the product is retained by the mineral under our standard elution conditions.

The 'feeding' protocol that we have outlined is a plausible model of prebiotic polymerization. The repeated incubations with low concentrations of activated monomers simulate conditions in environments where rocks are in constant contact with low concentrations of activated substrates or are episodically washed with higher concentrations. The model is robust in the sense that the presence of a large excess of unactivated substrate does not inhibit chain elongation. Thus a solution containing 1 µM of activated monomer in the presence of 100 µM unactivated monomer could yield only dimers in homogeneous solution but, on mineral surfaces, would extend a primer almost as efficiently as a pure 1 µM solution of the activated substrate.

The model as presented so far could lead to the covering of mineral surfaces by polymers of enormous length. In most relevant polymerizations, however, hydrolysis and chain termination would intervene. It is hard to discuss chain termination in a systematic way, as it depends on the idiosyncracies of the system under investigation and on the assumptions made about the prebiotic environment. The consequences of hydrolysis can be treated more systematically.

A simple calculation shows that the mean chain length $\bar{n} \approx \{[t_h(1/2)]/[t_c(1/2)]\}^{0.5}$ where $t_c(1/2)$ is the half-time for

chain elongation and $t_h(1/2)$ is the half-time for hydrolysis of each of the bonds in a polymer. It follows that, even if as much as one residue were added each day, the bonds in the resulting polymer would have to be stable for two or three years to achieve a mean chain length of 30. As elongation on the primitive Earth

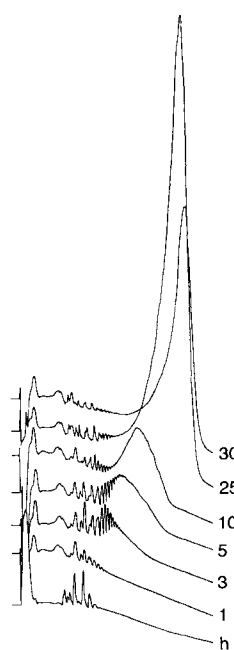


FIG. 4 Polymerization of aspartic acid on hydroxylapatite using *N*-ethyl-*N'*-dimethylaminopropyl-carbodiimide (EDAC) as condensing agent. The elution profile of a reaction mixture after a single incubation in the absence of mineral is labelled 'h'; the other traces correspond to products eluted after the designated numbers of cycles. Reactions were carried out at room temperature; each reaction tube contained 40 mg hydroxylapatite; the reaction solution contained 50 mM aspartic acid (pH 6.0) and 50 mM EDAC. To start the experiment, an aliquot (100 µl) of a freshly prepared solution of aspartic acid and EDAC is added to 40 mg hydroxylapatite in an 0.7 ml Eppendorf tube at room temperature. After 3.5 h, the tube is centrifuged, the supernatant is removed, and an aliquot of freshly prepared reaction mixture is added to begin the next cycle. After three cycles, the tubes are stored overnight at 4 °C in the absence of supernatant. After an appropriate number of cycles, the solid is washed 4 times with 250 µl water, and then eluted with each of two 250-µl aliquots of 20 mM K₄P₂O₇ for 15 minutes. An aliquot of the combined eluate (25 µl) is mixed with 2 ml starting buffer (pH 8) and analysed by HPLC on RPC-5; the reaction products are eluted from the column using a linear gradient of NaClO₄ (pH 8; 0–0.06 M, 80 minutes).

cannot have been rapid, only stable polymers such as peptides and nucleic acids could have accumulated from aqueous solution. Esters, for example, would have hydrolysed too quickly.

We have described experiments with negatively charged polymers, but the model is not restricted to these compounds. In principle it should apply, for example, to the synthesis of sulphhydryl-containing polymers on metal sulphides or to the synthesis of positively charged polymers on cation-exchanging minerals. We have extended the method to the synthesis of copolymers in an obvious way by incubating with a mixture of activated amino acids or by incubating successively with different activated amino acids.

The arguments presented above suggest that the minerals on the primitive Earth would have provided a 'library' of surfaces for the exploration of molecular evolution. Substrate-mineral combinations that lead to surface catalysis would be advantageous as sites for the development of a genetic system, but catalysis, although beneficial, is not essential. The feature lacking in our experimental systems is the ability of adsorbed molecules to replicate. In this regard, studies of template-directed reactions of nucleotides on surfaces^{3,8,9} and experiments related to those described above but involving template-directed ligation, are encouraging⁴. The possibility of simpler surface-bound replicating systems remains to be explored. □

Received 8 February; accepted 19 March 1996.

1. Szostak, J. W. & Ellington, A. D. in *The RNA World* (eds Gesteland, R. F. & Atkins, J. F.) 511–533 (Cold Spring Harbor Lab. Press, Cold Spring Harbor, NY, 1993).
2. Bernal, J. D. *The Physical Basis of Life* (Routledge & Kegan Paul, London, 1951).
3. Gibbs, D., Lohrmann, R. & Orgel, L. E. *J. molec. Evol.* **15**, 347–354 (1980).
4. Acevedo, O. L. & Orgel, L. E. *Nature* **321**, 790–792 (1986).
5. Ferris, J. P. & Ertem, G. *J. Am. chem. Soc.* **115**, 12270–12275 (1993).
6. Kawamura, K. & Ferris, J. P. *J. Am. chem. Soc.* **116**, 7564–7572 (1994).
7. Ehler, K. W. & Orgel, L. E. *Biochim. biophys. Acta* **434**, 233–243 (1976).
8. Schwartz, A. W. & Orgel, L. E. *J. molec. Evol.* **21**, 299–300 (1985).
9. Holm, N. G., Ertem, G. & Ferris, J. P. *Orig. Life Evol. Biosphere* **23**, 195–215 (1993).

ACKNOWLEDGEMENTS. We thank the American Colloid Co. for a gift of volclay, the mineral from which the Na⁺-montmorillonite used in this study was prepared, and G. Arrhenius for a gift of illite. We also thank S. Bailey for manuscript preparation. Research at RPI was supported by the US NSF, and at The Salk Institute by NSCORT/EXOBIOLOGY and NASA.

Rapid braincase evolution between *Panderichthys* and the earliest tetrapods

Per E. Ahlberg*, Jennifer A. Clack† & Ervins Lukševičs‡

*Department of Palaeontology, The Natural History Museum, Cromwell Road, London SW7 5BD, UK

†University Museum of Zoology, Downing Street, Cambridge CB2 3EJ, UK

‡Latvian Museum of Natural History, 4 Krīšjāņa Barona Street, Riga, LV1050, Latvia

THE panderichthyids (or elpistostegids) are the most tetrapod-like fishes that still retain paired fins rather than limbs. During the transition from fish to tetrapod, the braincase, previously subdivided by a joint, was remodelled into a solid structure¹. Here we present the complete braincase of the fish *Panderichthys rhombolepis*, a Middle Devonian¹ member of the tetrapod stemgroup^{2,3}. *Panderichthys* has an externally tetrapod-like skull², but we show that the braincase retained the intracranial joint, conforming wholly to the generalized pattern of lobe-finned fish, and sharing no obvious derived features with tetrapods. This places the braincase transformation between *Panderichthys* and the earliest tetrapods exemplified by *Acanthostega*⁴. The timing of the braincase transformation closely matches that of the limbs.

There are also striking similarities with the braincase transformation in the lungfish lineage. Both phenomena may reflect developmental linkages and canalization.

The braincase of *Panderichthys* (Figs 1, 2 and 3c, d) resembles the osteolepiform *Eusthenopteron*⁵ (Fig. 3a, b), a less crownward (more primitive) member of the tetrapod stem group, but also compares well with the living coelacanth *Latimeria*⁶. Like these taxa it has an intracranial joint, which runs between the basi-sphenoid and basioccipital ventrally, and immediately anterior to the otic capsule dorsally. Anteriorly, the basioccipital divides into left and right 'otic shelves' either side of a basicranial fenestra, which contains an arcual plate and would in life have accommodated the anteriormost part of the notochord. A lateral commissure (Fig. 2a, b) straddles the jugular canal and carries the hyomandibular articulations. The hyomandibula and basipterygoid process resemble those of *Eusthenopteron*. Two apparently unique characters of *Panderichthys* (or panderichthyids) are the vestibular fontanelle, which is rather small and anteriorly placed for a sarcopterygian (compare Fig. 3b, d, j), and the dorsally displaced optic tract, which reflects the position of the orbit. However, the general resemblance to *Eusthenopteron* is very close.

By contrast there are profound differences between the braincases of *Panderichthys* and the more crownward (less primitive) stem tetrapod *Acanthostega*⁴ (Fig. 3e, f). In the latter there is no lateral commissure, jugular groove, basicranial fenestra, arcual plate or intracranial joint. The anterior parts of the basioccipital and otic capsule have been drastically shortened, the hyomandibula has become a stapes with its footplate lodged in the enlarged and dorsally displaced vestibular fontanelle^{4,7}, and the basipterygoid process has changed shape.

This character distribution contrasts with the external morphology of the head. *Panderichthys* displays external tetrapod-like characters such as frontal bones, flattened snout, posteriorly

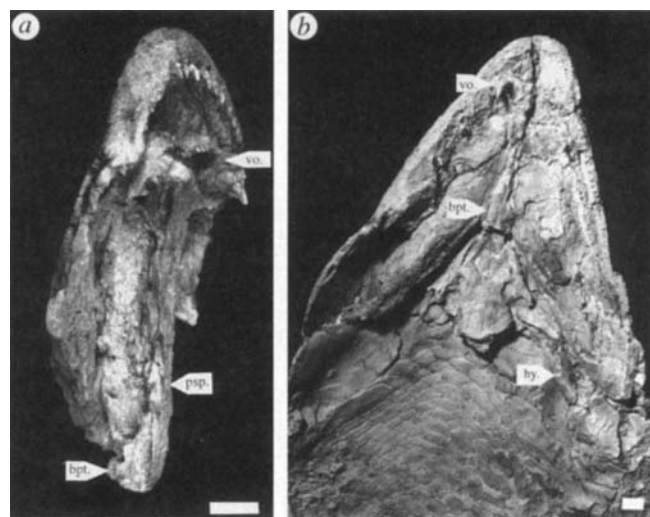


FIG. 1 a, Specimen LDM 43/4002, anterior (ethmosphenoid) portion of the braincase of *Panderichthys rhombolepis* together with partial skull roof, vomers (vo.) and parasphenoid (psp.). Note well-preserved optic nerve foramen (II) and basipterygoid process (bpt.). Right ventrolateral view. A left ventrolateral view of this specimen was presented by Worobjewa²⁸; we agree with her interpretation of the basipterygoid process and associated arterial foramina, but the 'nerve foramina' she identified appear to be artefacts of preparation. b, Specimen LDM 60/123, dorsoventrally flattened head and anterior part of the body in ventral view. Uniquely, this specimen preserves the whole braincase, as well as a crushed hyomandibular (hy.), but the anterior part is less well exposed than in 43/4002; it also includes the anterior end of the vertebral column, but this was removed before the photograph was taken. Scale bars, 10 mm. Both specimens were collected from Lode Quarry in Latvia (Gauja Formation, probably Upper Givetian¹) by L. Lyarskaya, and belong to the Latvian Museum of Natural History (Latvijas Dabas Muzejs) in Riga.

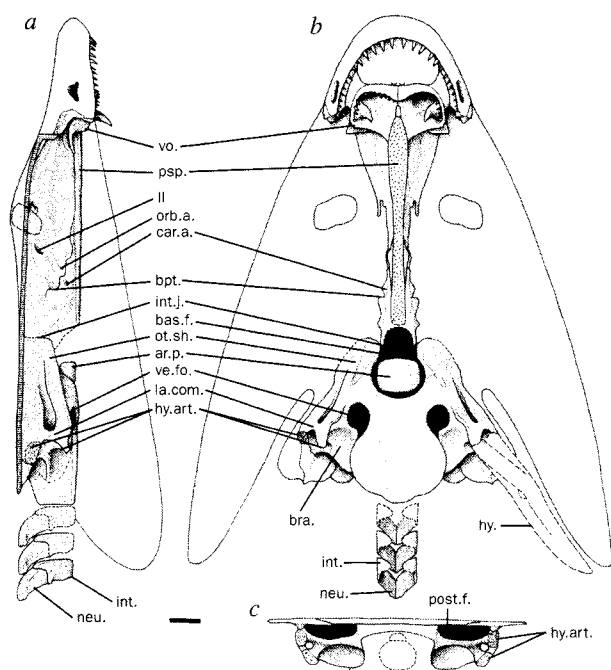
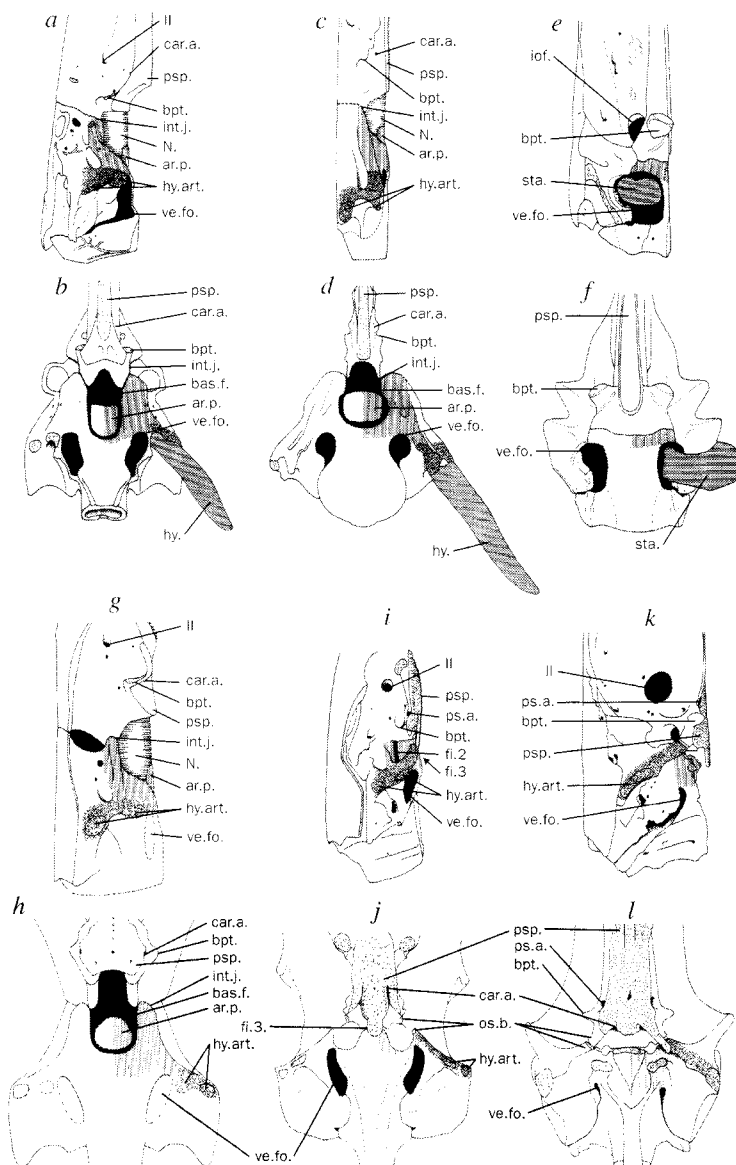


FIG. 2 a–c, Reconstructed braincase and three most anterior vertebrae of *P. rhombolepis* in lateral, ventral and posterior views, superimposed on outlines of skull. Dermal snout bones have been left in place; in a, horizontal hatching indicated cutaway view of skull roof. The braincase shows generalized sarcopterygian features; intracranial joint (int.j.), basicranial fenestra (bas.f.), 'otic shelves' (ot.sh.), arcual plate (ar.p.), and lateral commissures (la.com.) carrying the hyomandibular articulations (hy.art.). The dorsal part of the intracranial joint is obscured by the adjacent palatoquadrate. The basiterygoid process (bpt.) and the foramina for the carotid artery (car.a.) and artery of the orbital margin (orb.a.) resemble those of *Eusthenopteron*⁵, but the optic nerve foramen (ll) is displaced dorsally; other foramina could not be detected owing to crushing and poor visibility alongside palatoquadrate. The vestibular fontanelles (ve.fo) are small, and bounded posteriorly by concave areas (bra.) that may be branchial muscle attachments. The post-temporal fossae (post.f.) are very large. The hyomandibular (hy.) cannot be reconstructed in detail. Parts of four vertebrae are preserved in LDM 60/123, but the first and fourth are very incomplete. The neural arches (neu.) and intercentra (int.) are sutured together, but the contact differs from that shown in ref. 29 in trunk vertebrae. Both intercentra and neural arches are ossified as separate left and right halves. There are no pleurocentra, zygapophyses or obvious rib attachments. Scale bar, 10 mm.

FIG. 3 Posterior halves of braincases in lateral and ventral views. a–f, The tetrapod stem-group members *Eusthenopteron* (a, b; from ref. 5), *Panderichthys* (c, d) and *Acanthostega* (e, f). g–j, The lungfish stem-group members *Porolepis* (g, h); composite reconstruction from ref. 30) and *Youngolepis* (i, j; from ref. 22). k, l The primitive actinopterygian *Mimia* (from ref. 23). Not to scale. In the tetrapod stem group, *Eusthenopteron* and *Panderichthys* resemble each other much more than *Acanthostega*. In the two former genera the anterior part of the basioccipital region (vertical hatching), measured from the anterior margin of the vestibular fontanelle, is very long, and includes a basicranial fenestra with an arcual plate. This fenestra housed the notochord; in lateral view (a, c) the notochord (N.) has been reconstructed. In *Acanthostega* the anterior part of the basioccipital region is much shorter and lacks a fenestra. *Eusthenopteron* and *Panderichthys* both have long hyomandibulae (hy., horizontal hatching) carried by lateral commissures (stipple). *Acanthostega* has a short stapes (sta., horizontal hatching) and lacks the lateral commissure (position of footplate shown by horizontal hatching in e). It also has a very different basiterygoid process (bpt.), and a single large interorbital foramen (iof.), which probably admitted all the blood vessels that enter the hypophysial fossa. Note that the vestibular fontanelle of *Eusthenopteron* is shown as preserved (after Fig. 93 of ref. 5), not with a reconstructed cartilage rim (as in Fig. 86 of ref. 5). In the lungfish stem group, the poorly preserved *Porolepis* (g, h) clearly resembles *Eusthenopteron* and *Panderichthys*. This braincase structure represents the ancestral condition for both lungfish and tetrapod lineages (see Fig. 4). In *Youngolepis* (i, j) the intracranial joint has been lost, leaving only three fissures (fi. 1–3), corresponding respectively to the dorsal part of the joint, the lateral gap between the 'otic shelf' and arcual plate, and the junction between basioccipital and basisphenoid. The anterior part of the basioccipital is very short. The primitive actinopterygian *Mimia* (k, l) represents the sister group of the Sarcopterygii. Note the similarities with *Youngolepis*; the lateral commissure forms an 'otic-sphenoid bridge'²³ (os.b.) with a process for the basisphenoid, the carotid foramina are posterior rather than lateral to the parasphenoid (the lateral position is occupied by the afferent pseudobranchial artery, ps.a.), and the pro-otic and anterior basioccipital regions are short. The latter feature is also shared by *Acanthostega*.



open spiracular notches and dorsally placed orbits, and is thus much more derived than *Eusthenopteron*^{2,3,5,7}.

Panderichthys and *Acanthostega* are not successive plesions in the tetrapod stem group (Fig. 4). Unfortunately the braincases of *Elginerpeton*⁸ and *Ventastega*⁹, which probably intercalate between them, are unknown; however, *Ventastega* has an *Acanthostega*-like palate, which probably implies a similar braincase, and the more fragmentary *Elginerpeton* seems almost as derived⁸. We suggest that the braincase transformation occurred below the *Ventastega* node, and probably below the *Elginerpeton* node. It was thus probably quite an abrupt event. Below the *Panderichthys* node there had been no significant change in braincase construction since the origin of the Sarcopterygii (Fig. 4), despite considerable variation in skull shape. Above the *Acanthostega* node, the new braincase morphology persisted into the tetrapod crown group¹⁰, although the notochord gradually withdrew from the cranium and the parasphenoid extended posteriorly to cover the basioccipital.

The apparent rapidity of the braincase transformation contrasts oddly with most other aspects of the transition from fish to tetrapod, which began below the *Panderichthys* node and continued gradually up to the divergence of the amphibian and amniote lineages at the crown group node^{3,8,11} and beyond. Interestingly, only some parts of the braincase were affected. The occipital arches and opisthotics remained essentially unchanged in the tetrapod stem group and even into the amniote lineage, whereas the ethmoid and nasal capsules became unossified (probably a pedomorphic trait) but do not seem to have undergone marked structural change.

Only the pelvis and the distal parts of the limbs seem to show the same timing of character change as the braincase. The pectoral

fins of *Panderichthys*, like those of *Eusthenopteron*^{5,12}, lack digits, whereas *Acanthostega* has limbs with digits¹³ and a characteristic early-tetrapod pelvis¹⁴. *Elginerpeton*^{8,15} is associated with ilia and limb bones comparable to *Ichthyostega*, whereas *Ventastega*⁹ has a typical tetrapod ilium, clavicle and interclavicle; we predict that both genera will be found to possess digits. Although the pelvis of *Panderichthys* is unknown, it appears that the transformation of the limbs was rapid. As with the braincase, this sudden structural reorganization was preceded and followed by periods of very gradual change.

The braincase transformation seems to relate functionally to lengthening of the snout¹⁶, strengthening of the skull, and changes in the nature and range of palatal movement⁷, which were probably related to the apparently simultaneous loss of the bony gill cover. The origin of digits obviously modified locomotion, although not necessarily towards terrestrial walking¹³. The temporal correlation between these transformations is probably significant. On the one hand it suggests a change in the mode of life and associated selection pressures; on the other, it may point to the existence of developmental links between the two areas, so that some of the observed changes could be pleiotropic. The latter proposition should be testable in the light of developmental genetics.

Recent sarcopterygian phylogenies^{17–21} imply that tetrapods are not the only sarcopterygians to have lost the intracranial joint; it also disappeared, independently, in lungfishes (Fig. 4). The Lower Devonian stem-group lungfish *Youngolepis*²² resembles *Acanthostega* in lacking a basicranial fenestra and arcual plate, and in having the anterior parts of the otic capsule and basioccipital strongly foreshortened (Fig. 3i, j). These features of *Youngolepis*, as well as the relationships of the lateral commissure, basisphenoid and carotid artery foramina, are exact recreations of those in primitive actinopterygians^{23,24} (Fig. 3k, l). (The features are obscured in later lungfishes owing to the fusion of the palatoquadrate with the braincase.) In tetrapods the correspondence with actinopterygians is less precise, partly because of the loss of the lateral commissure. Outgroup comparisons with acanthodians^{4,25} suggest that the actinopterygian pattern corresponds essentially to the primitive osteichthyan condition. The cranial ontogeny of recent lungfishes and tetrapods compares very closely with that of both actinopterygians and chondrichthyans, which have never had intracranial joints²⁶.

The implication is that loss of the joint in lungfish and tetrapods in each case involved major character reversal, leading to retrieval of more primitive developmental pathways, and was not derived by terminal additions to an existing scheme of development.

Panderichthys shows that the tetrapod braincase structure evolved more abruptly than the external skull morphology. The transformation seems to have coincided with the origin of digits and tetrapod pelvis, at the internode traditionally taken to mark the 'origin of tetrapods'^{2–4,8,9,13,27}. Although the hypothesis will need to be tested through the investigation of other Devonian stem tetrapods, such as the enigmatic *Ichthyostega*⁵, we infer tentatively that this latter correlation is a real phenomenon reflecting functional and/or developmental linkage, rather than merely being an artefact of poor sampling. □

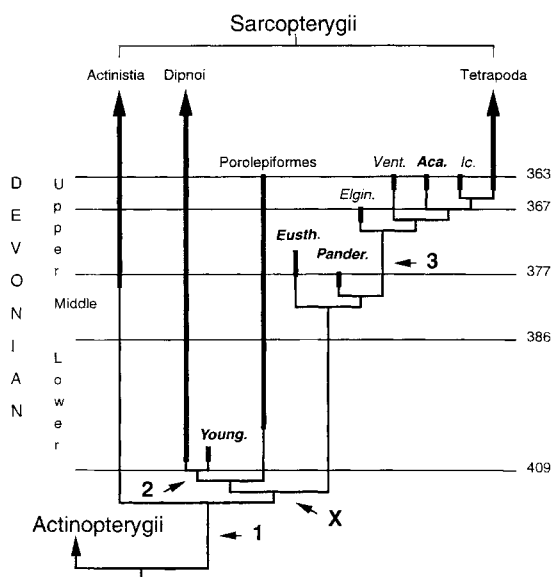


FIG. 4 Cladogram with timescale for a selection of sarcopterygians; topology based on refs 3, 8 and 21. Numbers on the right show age in millions of years. Lines with arrowheads indicate groups surviving to the present day. Genera labelled in bold are discussed in detail in the text. *Young*, *Youngolepis*; *Eusth.*, *Eusthenopteron*; *Pander*, *Panderichthys*; *Elgin*, *Elginerpeton*; *Vent*, *Ventastega*; *Aca.*, *Acanthostega*; and *Ic*, *Ichthyostega*. Following Lebedev and Coates¹¹, the tetrapod crown group (labelled Tetrapoda) is rooted in the Devonian. The node labelled X marks the beginning of the lungfish and tetrapod stem groups; the tetrapod stem group runs from this point up to the base of the Tetrapoda arrow. A traditional definition of Tetrapoda, based on the presence of limbs⁸, would include *Elginerpeton* and all the more crownward 'Devonian tetrapods'. *Eusthenopteron* has been used to represent the Osteolepiformes, an assemblage that may be either a monophyletic or a paraphyletic component of the tetrapod stem group. The numbers indicate: 1, the origin of the sarcopterygian jointed braincase; 2, the loss of the joint in the lungfish lineage; and 3, the loss of the joint in the tetrapod lineage.

Received 12 December 1995; accepted 14 March 1996.

1. Mark-Kurik, E. in *Abstracts of the Second Baltic Stratigraphic Conference* (eds Grigelis, A., Jankauskas, T.-R. & Mertinienė, R.) 57 (Baltic Stratigraphic Association, Vilnius, 1993).
2. Vorobyeva, E. I. & Schultze, H.-P. in *Origins of the Higher Groups of Tetrapods: Controversy and Consensus* (eds Schultze, H.-P. & Trueb, L.) 68–109 (Comstock, Ithaca, NY, 1991).
3. Ahlberg, P. E. & Milner, A. R. *Nature* **368**, 507–514 (1994).
4. Clack, J. A. *Nature* **369**, 392–394 (1994).
5. Jarvik, E. *Basic Structure and Evolution of Vertebrates Vol. 1* (Academic, London, 1981).
6. Millot, J. & Anthony, J. *Anatomie de Latimeria chalumnae. Tome I. Squelette, muscles et formations de soutien* (CNRS, Paris 1958).
7. Clack, J. A. *Nature* **342**, 425–430 (1989).
8. Ahlberg, P. E. *Nature* **373**, 420–425 (1995).
9. Ahlberg, P. E., Lukševičs, E. & Lebedev, O. *Phil Trans. R. Soc., Lond. B* **343**, 303–328 (1994).
10. Clack, J. A. *Zool. J. Linn. Soc.* (in the press).
11. Lebedev, O. A. & Coates, M. I. *Zool. J. Linn. Soc.* **114**, 307–348 (1995).
12. Vorobyeva, E. I. & Kuznetsov, A. in *Fossil Fishes as Living Animals* (ed. Mark-Kurik, E.) 131–140 (Academy of Sciences of Estonia, Tallinn, 1992).
13. Coates, M. I. & Clack, J. A. *Nature* **347**, 66–69 (1990).

14. Coates, M. I. & Clack, J. A. *Bull. Mus. natn. Hist. nat. Paris* **17**, 373–388 (1995).
15. Ahlberg, P. E. *Nature* **354**, 298–301 (1991).
16. Thomson, K. S. *Zool. J. Linn. Soc.* **46**, 223–253 (1967).
17. Maisey, J. G. *Cladistics* **2**, 201–256 (1986).
18. Panchen, A. L. & Smithson, T. R. *Biol. Rev.* **62**, 341–438 (1987).
19. Ahlberg, P. E. *Zool. J. Linn. Soc.* **103**, 241–287 (1991).
20. Chang, M.-M. & Smith, M. M. *J. vert. Paleontol.* **12**, 294–312.
21. Cloutier, R. & Ahlberg, P. E. *Géobios Mém. Spéc.* **19**, 241–248 (1995).
22. Chang, M.-M. *The Braincase of Youngolepis, a Lower Devonian Crossopterygian from Yunnan, South-Western China thesis*, Univ. Stockholm, 1982.
23. Gardiner, B. G. *Bull. Br. Mus. nat. Hist.* **37**, 173–428 (1984).
24. Ahlberg, P. E. *J. Morph.* **220**, 319 (1994).
25. Moy-Thomas, J. A. & Miles, R. S. *Palaeozoic Fishes* 2nd edn (Chapman & Hall, London, 1971).
26. De Beer, G. R. *Development of the Vertebrate Skull* (Oxford Univ. Press, 1937).
27. Vorobyeva, E. *The problem of the Terrestrial Vertebrate Origin* (Nauka, Moscow, 1992).
28. Worobjewa, E. J. *Palaeontographica A* **143**, 221–229 (1973).
29. Vorobyeva, E. I. & Tsessarsky, A. A. *J. gen. Biol.* **47**, 735–747 (1986) (in Russian).
30. Jarvik, E. *Meddr. Grönland* **187**(2), 1–307 (1972).

ACKNOWLEDGEMENTS. We thank H. Birnieks for taking the photographs for Fig. 1. P.A. thanks to the Natural History Museum for funding his visit to Latvia in 1995.

CORRESPONDENCE and requests for materials should be addressed to P.E.A. (e-mail: pea@nhm.ac.uk).

Role for heavy metals in forest decline indicated by phytochelatin measurements

James E. Gawel*, Beth A. Ahner†, Andrew J. Friedland‡ & François M. M. Morel†

* Civil and Environmental Engineering, 48-320, MIT, Cambridge, Massachusetts 02139, USA

† Geology Department, Princeton University, Princeton, New Jersey 08544, USA

‡ Environmental Studies Program, Dartmouth College, Hanover, New Hampshire 03755, USA

FOREST decline in the United States and Europe has been documented for a number of tree species^{1,2} and atmospheric pollutants from industrial sources, such as acids or oxidants, are thought to be partly responsible^{1–4}. Heavy metals have also been implicated because their deposition pattern is correlated with forest decline^{5–11}, but so far there has been no direct evidence for a physiological link between tree damage and exposure to metals. Here we use the concentrations of phytochelatins, which are intracellular metal-binding peptides that act as specific indicators of metal stress^{12,13}, to show that metals are indeed likely to be a contributing factor in the decline of forests in the northeastern United States. Phytochelatin concentrations in red spruce, a species in decline, are higher than in balsam fir, a species which is not. Concentrations increase with altitude, as does forest decline, and they also increase across the region in forest stands that show increasing levels of tree damage.

We collected current-year foliage from red spruce and balsam fir trees at 1,000 m on Whiteface Mountain in New York throughout the 1993 growing season to measure phytochelatin concentrations in the needles. Phytochelatins, intracellular metal complexing agents with the formula $(\gamma\text{-Glu-Cys})_n\text{-Gly}$, where $n = 2$ to 11, are produced specifically by plants as a response to sublethal concentrations of heavy metals^{14,15}. As can be seen from Fig. 1, mean foliar phytochelatin concentrations in red spruce were consistently higher than in balsam fir from June to August, with the largest and most significant ($P < 0.001$) difference occurring in mid-July at the peak of the growing season. In addition, the phytochelatin concentrations in red spruce in mid-July are significantly higher ($P < 0.001$) than at any other time sampled. (The relatively high standard deviations for the phytochelatin data reflect natural tree-to-tree variability, not analytical error; see Fig. 1 legend.) Balsam fir do not exhibit this peak, but rather stay at a consistently low level throughout the season. The

peak phytochelatin concentration measured in red spruce (0.5 nmol per g dry wt) is comparable to that reported for the leaves of sycamore maple growing on mine tailings (although the root concentrations in these trees were about 100 times higher — 4 and 500 nmol per g dry wt, respectively)¹³.

The number of standing dead red spruce trees increases sharply with elevation^{16,17}. If heavy metals are partly responsible for this trend, we would expect to see a similar increase in the phytochelatin concentrations in red spruce needles. Our measurements of current-year foliage from visually healthy red spruce trees confirm this expectation (Fig. 2). Phytochelatin concentrations, which again peaked in mid-July, were highest at 1,000 m and decreased systematically and significantly ($P < 0.001$) with decreasing elevation. No such trend was seen in balsam fir (data not shown).

As a further test of the relationship between heavy metals and the decline of forests in the northeastern United States, we sampled red spruce stands showing varying degrees of decline at 1,000 m elevation on nine mountains spanning New York, Vermont and New Hampshire. New foliage from five healthy trees was collected once a week from each location during July 1994 and all values for each site were pooled together. The percentage of red spruce trees that were standing dead was used as a measure of forest decline¹⁷. The data presented in Fig. 3 show a systematic and significant ($P < 0.001$) increase in phytochelatin concentrations that corresponds to the extent of tree damage in the stand. The highest concentrations of phytochelatins measured were from sites in the Green Mountains of Vermont and the

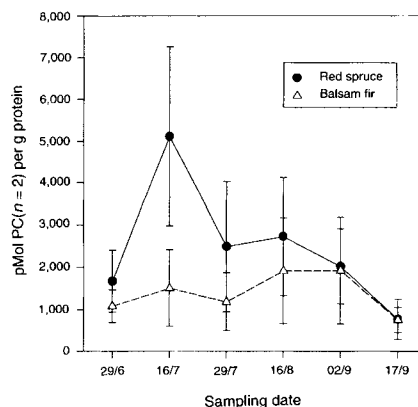


FIG. 1 Phytochelatin (pc) concentrations in current-year foliage from red spruce (*Picea rubens*) and balsam fir (*Abies balsamea*) in 1993 on the Whiteface Mountain massif, Adirondack Mountains, New York. Phytochelatin concentrations are given in picomoles (10^{-12}) of the $n = 2$ chain length (that is, $(\gamma\text{-Glu-Cys})_2\text{-Gly}$) normalized to grams of soluble protein, determined by BCA protein assay (Pierce Chemical) on needle extracts. Longer chain lengths of phytochelatins were not detected. Samples were collected from opposite sides of six red spruce and three balsam fir trees in the same stand at 1,000 m elevation using a pole pruner. Foliage from both aspects were pooled into one sample and stored immediately in liquid nitrogen before analysis. Samples were pulverized in liquid nitrogen in a ceramic mortar and pestle, and then homogenized on ice in 0.01 M methanesulphonic acid using a ground-glass pestle and tube (Glas-Col). The homogenate was centrifuged at 14,000 r.p.m. for 30 min at 4 °C and the supernatant refrigerated. Quantification of phytochelatin using reverse-phase HPLC with fluorometric detection is described in ref. 24. Error bars are calculated as standard deviations from the mean values. Replicate analyses of the same samples showed less than 10% relative error, whereas samples taken from different branches on the same tree and from different trees in the same stand produced up to twofold and fivefold differences, respectively. (Such variability is also found in metal concentrations measured in red spruce foliage⁹.) Statistical analyses of variation between tree species for a given sampling date were calculated using the unpaired Student's t -test; ANOVA analysis was used to determine the variation within a species over time.

Adirondack Mountains of New York, the areas most severely affected by forest decline⁷. These data strongly imply that metal stress is a cause, rather than an effect, of tree damage. Because phytochelatin concentrations were measured in visually healthy red spruce trees in each stand, they do not simply reflect the extent of individual tree damage. Rather, these high phytochelatin concentrations reveal that trees in the more affected stands, including the healthy ones, are under metal stress. This stress probably reflects increased exposure to heavy metals, although we cannot rule out the indirect effects of some other primary causes such as calcium or magnesium deficiency¹⁸.

The most parsimonious interpretation of all the available field data, which show consistency between the pattern of phytochelatin concentrations and the pattern in tree damage according to tree species, elevation and geographic distribution, is that heavy metals are a contributing cause of forest decline in the northeastern United States. A more direct study of the relationship between heavy metal exposure, phytochelatin production, and growth in red spruce is necessary to establish the degree of 'metal stress' indicated by our measurements of phytochelatin concentrations. A consensus is now emerging that multiple causative factors are involved in forest decline, probably in conjunction with freezing injury^{1,19}. Oxidant damage, recently shown to be a contributing factor in the reduced growth of loblolly pine in the southeastern United States²⁰, has also been implicated in red spruce decline in northeastern forests^{21,22}. It is known that glutathione depletion from the production of phytochelatin results in an increased sensitivity to oxidative damage in plants²³. It thus seems not unlikely to us that forest decline in the northeastern United States is the result of a synergistic effect of toxic metals and oxidants originating in air pollution³. In any case, elevated phytochelatin levels in red spruce provide *prima facie* evidence that trace-metal stress is contributing to their decline. Thus, abatement of atmospheric heavy-metal pollution should probably be considered as one of the elements of a policy directed at reversing forest decline in this region. Further, phytochelatin measurements in other forested regions of the world may allow us to determine whether the role of heavy metals in forest decline is a general phenomenon or limited to the northeastern United States. □

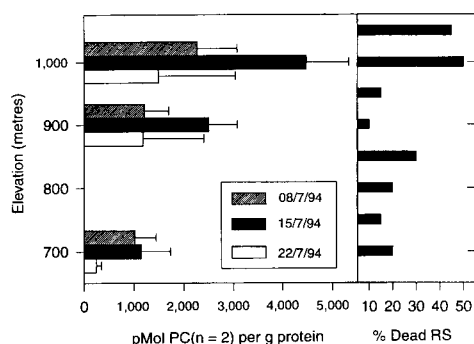


FIG. 2 Phytochelatin (PC) concentrations in new foliage from visually healthy red spruce (that is, showing less than 10% needle loss or discolouration) in 1994 along an elevational gradient (1,000, 900 and 700 m) on the Whiteface Mountain massif, Adirondack Mountains, New York, compared with per cent standing dead red spruce density. Dead red spruce as a per cent of total standing red spruce density was taken from ref. 16. Sample collection (from five red spruce trees at each elevation), extraction and analysis are described in Fig. 1 legend, with the substitution here of a Brinkmann Instruments' Polytron homogenizer for sample preparation. Similar results, as compared to those shown here and in Fig. 3, were obtained by normalizing to both fresh weight (0.05–0.15 nmol PC per g fresh wt) and dry weight (0.2–0.6 nmol PC per g dry wt). Error bars are calculated as standard deviations from the mean values. ANOVA analysis was used to determine the significance of phytochelatin variation over time and as a function of elevation.

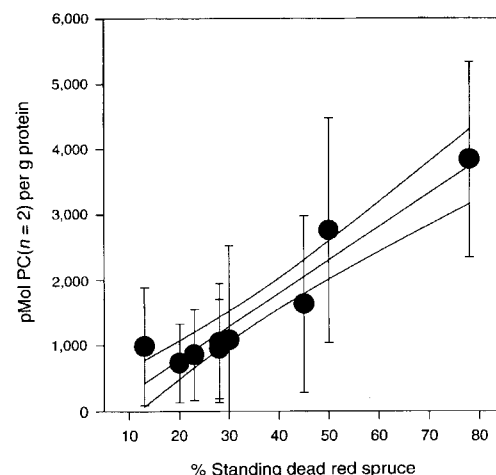


FIG. 3 Phytochelatin (PC) concentrations in new foliage from visually healthy red spruce in 1994 at 1,000 m on nine mountains in the northeastern United States, showing different degrees of forest damage. Sample collection (from five red spruce trees at each site), extraction and analysis are described for Fig. 2. Samples were taken once a week for four weeks in July 1994. All samples collected were pooled for one seasonal value at each site. The mountains sampled were, from left to right: Wildcat E, Monroe, Boott Spur and Jefferson (all in New Hampshire), Giant (New York), Clay (New Hampshire), Mansfield (Vermont), Whiteface (New York) and Abraham (Vermont). Per cent standing dead red spruce values were determined along 250 m transects following the 1,000 m elevation contour on all mountains, excluding Whiteface. Point-centred quarters were designated every 25 m, and the nearest standing red spruce stem larger than 10 cm in diameter in each quadrant was assessed as either dead or alive. Data for Whiteface Mountain were taken from ref. 16. Normalization to protein concentrations, dry weight and fresh weight all produced similar results. Error bars are calculated as standard deviations from the mean values for each mountain. ANOVA analysis was used to determine the significance of the correlation between phytochelatin concentrations and per cent standing dead red spruce for all samples, and the 99% confidence intervals are shown.

Received 12 April 1995; accepted 13 March 1996.

1. Eager, C. & Adams, M. B. (eds) *Ecology and Decline of Red Spruce in the Eastern United States* (Springer, New York, 1992).
2. Schulze, E.-D. *Science* **244**, 776–783 (1989).
3. Taylor, G. E. Jr, Johnson, D. W. & Andersen, C. P. *Ecol. Appl.* **4**, 662–689 (1994).
4. McLaughlin, S. B. J. *Air Pollut. Control Ass.* **35**, 512–534 (1985).
5. Johnson, A. H. & Siccama, T. G. *Environ. Sci. Technol.* **17**, 294A–305A (1983).
6. Goldberg, D. L. & Hüttermann, A. *Environ. Pollut. Ser. A* **38**, 375–381 (1985).
7. Herick, G. T. & Friedland, A. J. *Water Air Soil Pollut.* **53**, 151–157 (1990).
8. Friedland, A. J., Johnson, A. H. & Siccama, T. G. *Water Air Soil Pollut.* **29**, 233–243 (1986).
9. Friedland, A. J. & Johnson, A. H. *J. envir. Qual.* **14**, 332–336 (1985).
10. Smith, W. H. & Siccama, T. G. *J. envir. Qual.* **10**, 323–333 (1981).
11. Johnson, A. H., Siccama, T. G. & Friedland, A. J. *J. envir. Qual.* **11**, 577–580 (1982).
12. Schat, H. & Kalf, M. M. A. *Plant Physiol.* **99**, 1475–1480 (1992).
13. Grill, E., Winnacker, E.-L. & Zenk, M. H. *Experientia* **44**, 539–540 (1988).
14. Gekeler, W., Grill, E., Winnacker, E.-L. & Zenk, M. H. *Z. Naturf.* **44c**, 361–369 (1989).
15. Loeffler, S., Hochberger, A., Grill, E., Winnacker, E.-L. & Zenk, M. H. *FEBS Lett.* **258**, 42–46 (1989).
16. Battles, J. J., Johnson, A. H., Siccama, T. G., Friedland, A. J. & Miller, E. K. *Bull. Torrey Bot. Cl.* **119**, 418–430 (1992).
17. Craig, B. W. & Friedland, A. J. *Environ. Monit. Assess.* **18**, 129–143 (1991).
18. McLaughlin, S. B., Tjoelker, M. G. & Roy, W. K. *Can. J. For. Res.* **23**, 380–386 (1993).
19. Sheppard, L. J. *New Phyt.* **127**, 69–82 (1994).
20. McLaughlin, S. B. & Downing, D. J. *Nature* **374**, 252–254 (1995).
21. Madamanchi, N. R., Hausladen, A., Aischer, R. G., Amundson, R. G. & Fellows, S. *New Phyt.* **118**, 331–338 (1991).
22. Vann, D. R., Strimbeck, G. R. & Johnson, A. H. *For. Ecol. Mgt* **51**, 69–79 (1992).
23. De Vos, C. H. R., Vonk, M. J., Vooijs, R. & Schat, H. *Plant Physiol.* **98**, 853–858 (1992).
24. Ahner, B. A., Price, N. M. & Morel, F. M. M. *Proc. natn. Acad. Sci. U.S.A.* **91**, 8433–8436 (1994).

ACKNOWLEDGEMENTS. We thank A. Foster, C. White, B. Spranz, D. Lockwood and M. Roberts for their help in collecting samples; J. Reinfelder for his help in editing this manuscript; and the National Institute of Environmental Health Sciences, the NSF, the EPA, the Andrew W. Mellon Foundation, and the US Forest Service Northern Stations Global Change Research Program for their support.

CORRESPONDENCE and requests for materials should be addressed to J.E.G. (e-mail: waterboy@mit.edu).

Enhancement of selective listening by illusory mislocation of speech sounds due to lip-reading

Jon Driver

Department of Psychology, Birkbeck College, University of London, Malet Street, London WC1E 7HX, UK

MECHANISMS of human attention allow selective processing of just the relevant events among the many stimuli bombarding our senses¹. Most laboratory studies examine attention within just a single sense, but in the real world many important events are specified multimodally, as in verbal communication. Speech comprises visual lip movements as well as sounds, and lip-reading contributes to speech perception, even for listeners with good hearing, by a process of audiovisual integration². Such examples raise the problem of how we coordinate our spatial attention across the sensory modalities, to select sights and sounds from a common source for further processing. Here we show that this problem is alleviated by allowing some cross-modal matching before attentional selection is completed. Cross-modal matching can lead to an illusion, whereby sounds are mislocated at their apparent visual source³; this crossmodal illusion can enhance selective spatial attention to speech sounds.

The textbook case of selective attention involves listening to one conversation while ignoring others at a noisy party. Although many sounds enter the ear concurrently, we only have a full awareness for those that receive attentive processing⁴. Spatial mechanisms of auditory attention, and of visual attention, have been extensively studied in simplified laboratory tasks, and their neural substrates have recently been investigated with various neuroimaging methods¹. However, such laboratory studies examine spatial attention within just a single modality at a time, whereas many real-world attentional tasks are multimodal. For instance, selective listening usually involves lip-reading (the visual processing of lip movements and other cues given by seeing the speaker's tongue and teeth) in addition to auditory processing⁵.

Lip-reading when listening to a single auditory message gives rise to several dramatic cross-modal interactions. The McGurk effect⁶ arises when a speech sound is dubbed onto the soundtrack of a film that shows lip movements for a different but related sound. People typically hear a speech sound that is a compromise between the auditory and visual information, illustrating that visual processing can influence auditory experience^{2,7}. A second effect, the ventriloquist illusion³, arises when people mislocate sounds towards their apparent visual source. This applies powerfully for speech sounds and matching lip movements, as when the spoken soundtrack of a film seems to emanate from actors appearing on the cinema screen, despite a true sound source elsewhere.

The ventriloquist illusion has been well documented when single auditory messages are presented^{3,8,9}. However, its possible implications for auditory spatial attention, when multiple auditory messages are presented, have never been addressed. Here we test whether attention gets directed to the illusory position of the relevant speech sounds in such cases, as determined visually by their matching lip movements. If so, the ventriloquist illusion might actually enhance veridical selective listening. This would require that some cross-modal matching takes place before the operation of spatial attention is completed, as explained below.

Experiment 1 required subjects to repeat one verbal message while ignoring another. Both messages were spoken by the same voice at the same time, and both comprised a 'nonsense' sequence of three unrelated words. Unknown to our subjects, both messages also came from a common loudspeaker in mono. Thus no auditory cues indicated which sequences of words were targets, and which

distractors. This was indicated only by a video screen, which gave simultaneous visual lip-read information for just the target message (Fig. 1).

These lip movements were either shown on the side of the true mono sound source (Fig. 1a) or displaced at the location of a dummy loudspeaker on the other side (Fig. 1b). The ventriloquist illusion should scarcely affect the same-side condition, as the lips

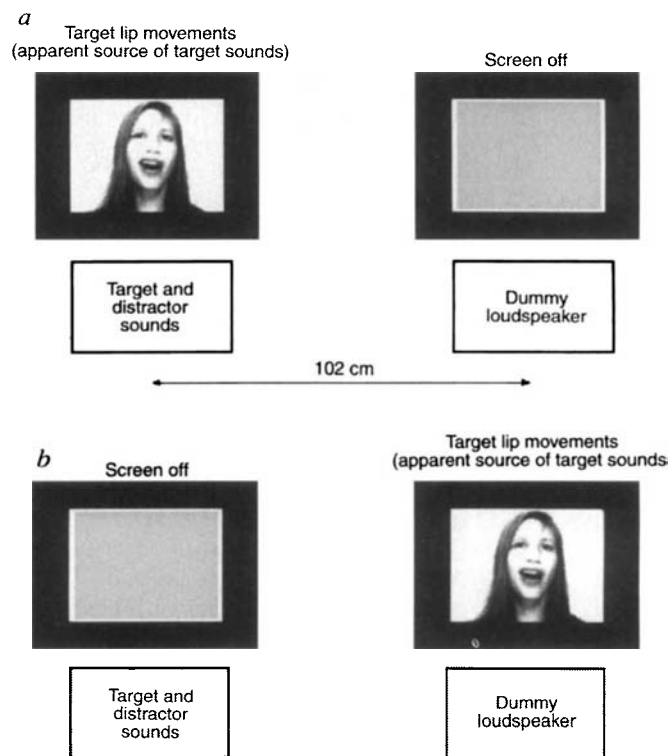


FIG. 1 The layout of apparatus is shown schematically, to scale, for both conditions of experiment 1: *a*, same-side condition; and *b*, displaced condition. The task was always to repeat the words spoken by the person appearing on-screen in a frontal full-face view. Each subject fixated the lip movements throughout, as the task was otherwise impossible (it was also impossible without audition). The subject's head was fixed to face directly between the two loudspeakers and monitors, which lay on a plane 198 cm in front of the subject. At any one time, only one monitor presented stimuli (lip movements for just the target words), and only one loudspeaker played sounds (both target and distractor words in mono). Which loudspeaker was active varied across subjects (the left is shown active in the illustrations). Within participants, the two monitors were used equally often on different blocks of trials, so that the lips could appear on the same side as the active loudspeaker (as in *a*) or on the other side near the dummy loudspeaker (*b*). The materials were video recordings of a woman pronouncing two-syllable, stress-initial words in synchrony with clicks that were presented to her over headphones every 500 ms, with pitch changes allowing her to anticipate when to pronounce a triplet. These clicks were not audible on the video recording for the sections presented. Two videotapes were made, each of 336 different words arranged into 112 triplets. No word appeared on both tapes. Each triplet had approximately 1 word every 500 ms. Concurrent target and distractor triplets were followed by 4.5 s of silence, during which participants had to repeat just the target triplet. At the presentation rate of 500 ms per word, inadvertent coarticulation cues might conceivably provide some auditory information about which sounds belong to a common triplet; however, this could not signal which words were targets (indeed, the task was impossible without vision). Moreover, any coarticulation cues would apply equally for the same-side and displaced conditions, and likewise for any possible contributions from auditory echoic memory. Target and distractor tapes were played in synchrony, by a video editor which aligned clicks at their start. Concurrent word pairs were roughly equated for spoken frequency; full details of these words are given elsewhere¹⁰. The synchronous auditory messages produced 46–48 dB(*b*) at the participant's location; 38 dB(*b*) was due to mains hum and fan noise from the apparatus.

were very close to all of the speech sounds. However, in the displaced condition (Fig. 1b), ventriloquism might pull the target sounds away from the distractor sounds, towards the side of the dummy loudspeaker where the lips appeared. Only the target sounds match these lip movements, so they might migrate towards the lips alone, leaving the distractor sounds behind. This should produce an illusory spatial separation between the sounds. We tested whether this potential illusion could enhance veridical selective listening, as previously found when true spatial separation is physically present within audition⁴.

The results (Fig. 2) show that selective report of the target message was indeed considerably more efficient in the displaced condition. We attribute this to an illusory lateral separation of the sounds, produced by ventriloquism for the target message (which all subjects experienced). Previous findings with the same verbal materials discount explanations that do not invoke ventriloquism¹⁰, as selective listening is usually worse rather than better when visual fixation is directed away from the physical location of target sounds, as in the present displaced condition. Moreover, the direction of fixation applied equally to the target and the distractor sounds in experiment 1, so this alone could not have enhanced perception for only the target sounds. Our results must therefore be due to the ventriloquist illusion altering the perceived location of just the target sounds. This provides a unique case of an illusion actually aiding veridical perception. Evidently, attention can be usefully focused on the virtual rather than actual location of a sound, as illusorily determined by vision.

In experiment 1, audiovisual integration could only improve performance. Experiment 2 was devised so that cross-modal

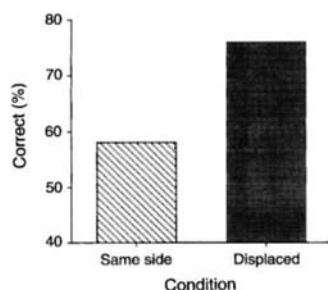


FIG. 2 Results from 12 naive participants in experiment 1, with a mean age of 22. The condition (same-side versus displaced lips) was changed every 14 triplets simply by turning one monitor off and the other on. The tapes were run through twice. On the second run, subjects repeated words that had previously been distractors (that is, the other tape now went to the active video screen). Across participants, particular words were counter-balanced over the same-side and displaced conditions, as was the sequence of these conditions. Responses in the selective repetition task were scored as exactly correct, incorrect or intrusions (repetitions of words from the distractor message). The mean percentages of correct reports are shown. The displaced condition showed better performance than the same-side condition ($T = 3$, $P < 0.01$ in a Wilcoxon matched-pairs signed-rank test). There were only 32 intrusions among the 4,032 responses in the displaced condition, and 92 intrusions for the 4,032 responses in the same-side condition. Although these intrusions were rare (less than 2% of responses overall), they arose significantly more often in the same-side condition ($T = 0$, $P < 0.01$), corroborating its inferiority. On debriefing, all subjects reported that the target sounds consistently appeared to come from wherever the lips were presented, thus demonstrating the ventriloquist illusion^{3,7}. Only one subject became unsure of this when pressed. Our experiment thus shows for the first time that ventriloquism still arises when multiple competing speech sounds are presented concurrently. More importantly, the results also reveal that ventriloquism does not merely affect phenomenology, but also the objective efficiency of veridical selective listening. In principle, the illusion might therefore be taken advantage of to enhance video communication. Mere spatial separation between a screen showing a person's lip movements and a loudspeaker relaying their corresponding speech sounds should lead to an illusory spatial separation of the speech sounds from any noise in the auditory transmission. As a result, the perceiver should be better able to ignore transmission noise, without requiring the video device itself to filter out this noise.

integration could now only hinder performance, to determine whether its influence on spatial selection arises even when unwanted. The target and distractor sounds were now spatially separated, appearing on different sides of a video screen. Thus the target message in this study could be selected merely by its spatial location within audition (Fig. 3a). Any lip movements were now presented only for the distractor message (Fig. 3b).

Performance was impaired by these distractor lip movements (Fig. 3c). This accords with the proposal² that cross-modal

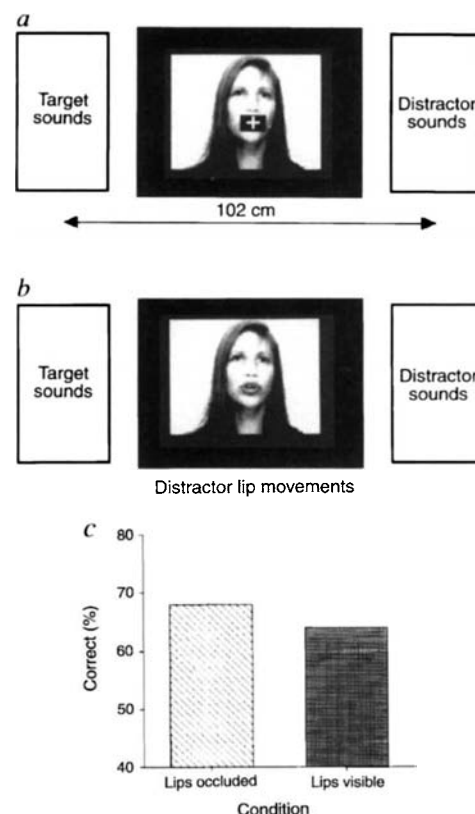


FIG. 3 The schematic layout is shown to scale for the two conditions of experiment 2 (a versus b). The apparatus lay on a plane 152 cm in front of the subject. Materials were as before. Target and distractor sounds now came from different loudspeakers, so the relevant message was specified by its auditory location (36° from the distractor sounds). The single central monitor always showed the recording of the distractor message, with lip movements either visible or occluded by a strip. Participants fixated either the centre of the occluding strip (a) or the distractor lip movements when these were visible at the same location (b). Adherence to the central fixation instructions was confirmed throughout by the experimenter. Every 14 triplets, the target and distractor sides were reversed in audition (and the participant instructed accordingly), by changing which recording went to the central monitor (in the illustrations, the right sounds match the lips on the monitor). The lip-read condition (occluded versus visible) was changed every 28 triplets. Across participants, particular words were counter-balanced over conditions, as was the sequence of these conditions. We anticipated that any involuntary audiovisual integration would impair performance when the lips were visible, for two possible reasons: (1) any ventriloquism would shift the distractor sounds towards the central matching lips, and thus closer to the target sounds; and (2) the lips provide further information about the distractor sounds which might then overwhelm the target sounds. Because either source of impairment would demonstrate that audio-visual integration between distractor stimuli can influence the efficiency of attentional selection, we did not seek to distinguish them. There were 12 new participants, with a mean age of 20. c, Mean percentages of correct reports, for the lips-visible and lips-occluded conditions. There were more correct reports with the lips occluded ($T = 1.5$, $P < 0.01$). From a total of 8,064 responses, there were just 49 intrusions with lips occluded and 79 with lips visible (again less than 2% overall). The trend for more intrusions with lips visible was unreliable ($T = 17.5$, N.S.).

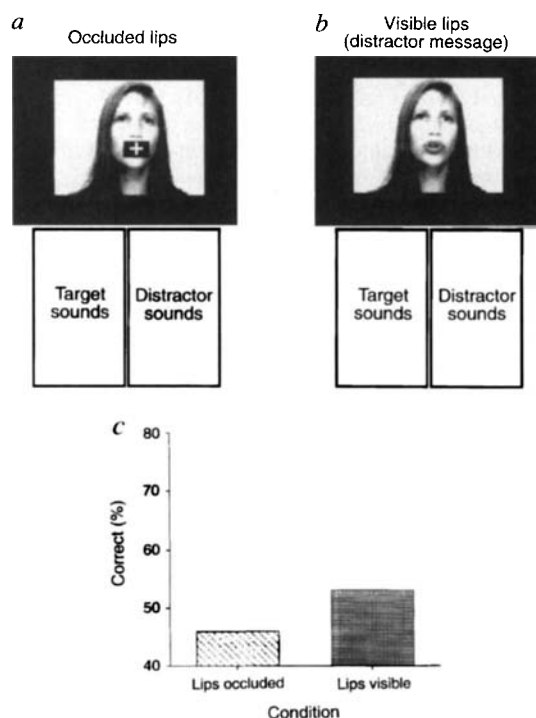


FIG. 4 The schematic layout is shown for the two conditions of experiment 3 (a versus b). The materials and procedure were as for experiment 2, except that the two speakers were now as close as possible. The target message was still specified by its auditory location, but the auditory separation of target and distractor messages was reduced to one-fifth of that used previously. Under these circumstances of difficult auditory segmentation, it was hypothesized that the potential segmentation benefits from the distractor lips should become more apparent. As before, the distractor lips were either occluded (a) or visible (b). There were 12 new subjects, with a mean age of 24. c, Mean percentages of correct reports for the lips-visible and lips-occluded conditions. Overall performance was worse than experiment 2 (compare c with Fig. 3c) owing to the reduced auditory spatial separation between target and distractor messages. More importantly, correct reports were now more frequent when the distractor lips were visible rather than occluded ($T = 1$, $P < 0.01$), reversing the outcome of experiment 2. From a total of 8,064 responses, there were just 62 intrusions with the lips visible, and 117 with lips occluded. Although these intrusions were rare (less than 3% of responses overall), they arose more often in the lips-occluded condition ($T = 10$, $P < 0.02$).

integration in speech perception arises automatically, even when unwanted. This proposal was based on findings^{2,11} that the McGurk effect is unchanged by instructions emphasizing one modality over the other. Our studies extend those findings in several respects. First, the previous studies^{2,11} only presented a single auditory event together with a single visual event, with at least one of these events always being an attended target. Hence experiment 2 provides the first demonstration of unintended cross-modal integration between auditory and visual events that are both distractors, in a situation where concurrent targets and distractors compete for attention within audition. A further difference is that the single auditory and visual events always had a common location in the studies examining the McGurk effect's susceptibility to intentions^{2,11}. Thus there was no opportunity for spatial selection in those studies, within or between modalities. Hence the previous studies have none of the implications for spatial attention that were addressed by our ventriloquism manipulation in experiment 1, and in our final study.

Experiment 3 was motivated by the following observation. Although reliable, the cost from distractor lip movements in experiment 2 was slight compared with the benefits that have previously been observed from adding target lip movements⁵. Moreover, several participants felt that the distractor lips actually helped them to separate target and distractor sounds, despite impairing performance overall. This suggests that the distractor lips may have induced two opposing effects. First, they impair performance (as shown in the overall result) by adding information that makes the auditory distractor words more potent. A second influence from the distractor lips may have mitigated this impairment, by aiding selection as follows. Cross-modal matching should group just the distractor sounds together with their corresponding lip movements. This would segment the distractor

sounds away from the target sounds, and thus should enable selection of the relevant message.

Any such beneficial influence should become more apparent when segmentation becomes the critical factor that limits performance. Experiment 3 was like experiment 2, but with a drastically reduced spatial separation between target and distractor sounds, so that selection could scarcely be based on audition alone (Fig. 4). Here the possible segmentation benefit from the distractor lips might outweigh the cost in terms of adding distractor information. Indeed, selective listening now became more efficient when the distractor lips were shown, reversing the result of experiment 2. This provides the first demonstration that adding information about distractors, within another modality, can sometimes make them easier to ignore.

Our studies also yield evidence that cross-modal matching can arise before auditory spatial selection is completed. Experiments 1 and 3 found that ventriloquism can enhance the efficiency of auditory selection, through the effects of cross-modal matching on the apparent location of sounds. This cross-modal matching must therefore arise before spatial auditory selection is completed. Experiment 2 found that cross-modal interactions still take place even when detrimental to auditory selection. However, in the real world their effects should usually be beneficial, because auditory localization is considerably poorer than visual localization¹². Audiovisual matching can compensate for this poor auditory acuity by locating sounds at their sources as coded with the spatial precision of vision. However, this cross-modal mechanism for enhancing auditory localization could not possibly help us to listen selectively if it arose only after auditory spatial selection was completed. Our findings reveal that in fact it arises before selection is completed, and as a result can enhance selective listening for target sounds that would otherwise be indistinct in location from distractor sounds. □

Received 21 February; accepted 21 March 1996.

1. Naatanen, R. *Attention and Brain Function* (Erlbaum, Hillsdale, NJ, 1992).
2. Massaro, D. W. *Speech Perception by Ear and by Eye* (Erlbaum, Hillsdale, NJ, 1987).
3. Bertelson, P. & Radeau, M. *Percept. Psychophys.* **19**, 531–535 (1976).
4. Cherry, E. C. *J. acoust. Soc. Am.* **25**, 975–979 (1953).
5. Reisberg, D. *Acta psychol.* **42**, 331–341 (1978).
6. McGurk, H. & MacDonald, J. *Nature* **263**, 747–748 (1976).
7. Massaro, D. W. & Cohen, M. M. *Curr. Dir. psychol. Sci.* **4**, 104–109 (1995).
8. Radeau, M. *Curr. Psychol. Cogn.* **13**, 3–51 (1994).
9. Witkin, H. A., Wapner, S. & Leventhal, T. *J. exp. Psychol.* **43**, 58–67 (1952).

10. Driver, J. & Spence, C. J. in *Attention and Performance Vol XV* (eds Umiltà, C. & Moscovitch, M.) 311–332 (Erlbaum, Hillsdale, NJ, 1994).
11. Massaro, D. W. in *Perspectives on Perception and Action* (eds Heuer, H. & Sanders, A. F.) 273–299 (Erlbaum, Hillsdale, NJ, 1987).
12. Blauert, J. *Spatial Hearing* (MIT Press, Cambridge, MA, 1983).

ACKNOWLEDGEMENTS. I thank A. Allport, G. Davis and M. Posner. This work was supported by the Medical Research Council (UK).

CORRESPONDENCE and requests for materials should be addressed to the author (e-mail: j.driver@psyc.bbk.ac.uk).

Central reorganization of sensory pathways following peripheral nerve regeneration in fetal monkeys

Sherre L. Florence*, Neeraj Jain*,
Marcie W. Pospichal*, Pam D. Beck*,
D. Lewis Sly† & Jon H. Kaas*

Departments of * Psychology and † Pathology, Vanderbilt University, Nashville, Tennessee 37240, USA

TRANSECTION of a sensory nerve in adults results in profound abnormalities in sensory perception, even if the severed nerve is surgically repaired to facilitate accurate nerve regeneration. In marked contrast, fewer perceptual errors follow nerve transection and surgical repair in children¹⁻³. The basis for this superior recovery in children was unknown. Here we show that there is little or no topographic order in the median nerve to the hand after median nerve section and surgical repair in immature macaque monkeys. Remarkably, however, in the same animals the representation of the reinnervated hand in primary somatosensory cortex (area 3b) is quite orderly. This indicates that there are mechanisms in the developing brain that can create cortical topography, despite disordered sensory inputs. Presumably the superior recovery of perceptual abilities after peripheral nerve transection in children depends on this restoration of somatotopy in the central sensory maps.

We assessed the effects of median nerve transection, surgical repair and regeneration in four immature macaque monkeys; in two monkeys the nerve was transected prenatally, and in the other two animals, the nerve was cut after birth. At 10–18 months of age, the distribution of peripheral nerve terminations in the cuneate nucleus and spinal cord from one of the fingers reinnervated by the regenerated nerve was labelled by injecting small amounts of horseradish peroxidase-conjugated tracers, and the topographic organization of the reinnervated hand representation in cortical area 3b was assessed using standard microelectrode mapping techniques. Results are compared to those of our earlier study of nerve cut and repair in adult macaques⁴, which had comparably long survival times (7–13 months) after the nerve section.

Injections into the reinnervated digit produced a much larger focus of label in the cuneate nucleus than the projection from the same digit on the contralateral, control hand (Fig. 1). Label transported from the reinnervated digit occupied nearly the entire median nerve innervation territory (the lateral half) of the pars rotunda in the cuneate nucleus, extending throughout the region where inputs from digits 1–3 terminate. Injections of the reinnervated digits also labelled much more extensive regions in the dorsal horn of the spinal cord than the matched injections in the opposite hand (not shown). Most important, the extent of label in both structures appeared to be similarly widespread, regardless of the age at which the nerve was cut (Fig. 1), and was comparable to that obtained after nerve cut and regeneration in adult macaques⁴. The widespread label probably reflects regeneration errors so that some of the neurons that normally innervate other parts of the median nerve territory regenerated into the injected finger⁴; however, new growth of the afferent inputs also may have occurred centrally^{5,6}. As a result of the abnormal distribution of inputs, information about neighbour relationships on the hand would be profoundly degraded in the cuneate nucleus and the dorsal horn of the spinal cord.

In marked contrast to the abnormal distribution of the peripheral nerve terminations in the cuneate nucleus and spinal cord after median nerve cut in immature monkeys, the somatotopy and the response properties of neurons in cortical area 3b were

remarkably normal. Receptive fields, the skin regions within which light cutaneous stimulation best activated the neurons, tended to be small and had somatotopically appropriate progressions in area 3b contralateral to the deprived hand. Figure 2 shows the sizes and somatotopic distribution of receptive fields in an animal that had median nerve cut prenatally (Fig. 2A, C) and in an animal that had an early postnatal nerve cut (Fig. 2B, D). Regardless of the age at which the nerve was cut, proximal phalanges innervated by the regenerated nerve were represented more superficially in the central sulcus, where area 3b is located, and the middle and distal phalanges were represented progressively deeper in the sulcus (see receptive fields 1–6 for penetration *c* in Fig. 2B). These features are characteristic of area 3b in normal monkeys⁷. However, there were local abnormalities that were variable across animals. Some receptive field progressions had abrupt reversals so that, for example, receptive field locations might shift from the proximal to distal hand and back (Fig. 2B,

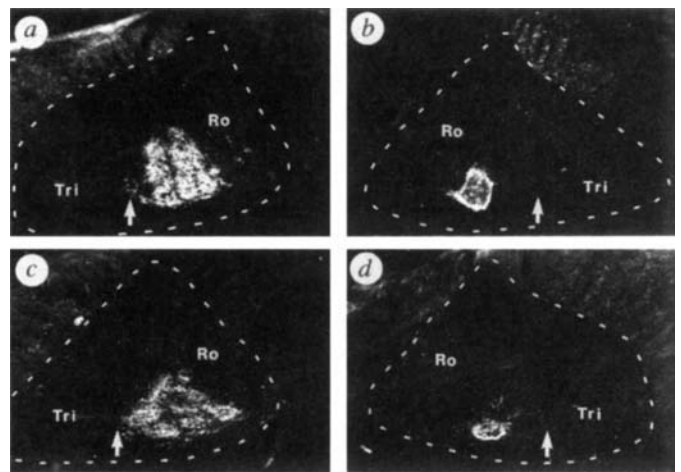


FIG. 1 The distribution of label in frontal sections through the pars rotunda of the cuneate nucleus following injections of cholera toxin subunit B conjugated to horseradish peroxidase (HRP) in digit 2 (D2) after median nerve cut, repair and regeneration at about 115 days of gestation (a) and on postnatal day 10 (c); the distribution of label from injections of D2 on the control hand is shown in b and d. Label appears as white in these dark-field photomicrographs. Dashed white lines indicate the borders of the cuneate nucleus; arrows indicate the lateral border of the pars rotunda (Ro). Dorsal is to the top and medial is towards the middle of the figure. Tri, pars triangularis of the cuneate nucleus.

METHODS. The median nerve was transected and rejoined in 2 fetal macaque monkeys (*Macaca mulatta*), estimated to be at 110 (case 92-74) and 115 (case 93-18) days gestation (normal gestation is about 165 days), and in 2 postnatal macaques, 10 and 14 days after birth. Fetal surgery was accomplished after the pregnant monkey was deeply anaesthetized with Isoflurane, and given a smooth-muscle relaxant (Brethine, 0.01 mg kg⁻¹, by intramuscular injection (i.m.)). A midline incision was made in the abdominal wall and the uterus exposed. A small incision was made in the uterus at a non-placental location, and stabilized with a purse-string suture. One arm of the fetus was located by palpation, then extracted from the uterine cavity into an amniotic pouch. The amniotic sac was incised to expose the fetal forearm, the median nerve dissected midway between the wrist and elbow and transected with surgical scissors. The cut ends of the nerve were then rejoined using 10.0 surgical silk, without attempting to maintain nerve alignment. The incisions were closed with suture and the anaesthesia was withdrawn. Antibiotics (penicillin, 20,000 units per lb body weight, i.m.), analgesics (Banamine, 1 mg kg⁻¹, i.m.), and a smooth-muscle relaxant (Brethine, 1.25 mg three times a day) were given postoperatively. The median nerve was transected in the postnatal monkeys as described. All monkeys survived for 10–19 months before terminal experiments were performed. Three to five days before terminal mapping experiments, small (10–12 μ l) subcutaneous injections of cholera toxin subunit B conjugated to HRP (1.0–0.5%) were made into the distal phalange of D2 bilaterally. Procedures for injection, killing and tissue processing have been described⁴.

penetration d). As shown in Fig. 2C and D, these reversals did not disrupt the overall topographic progressions within the map of the hand.

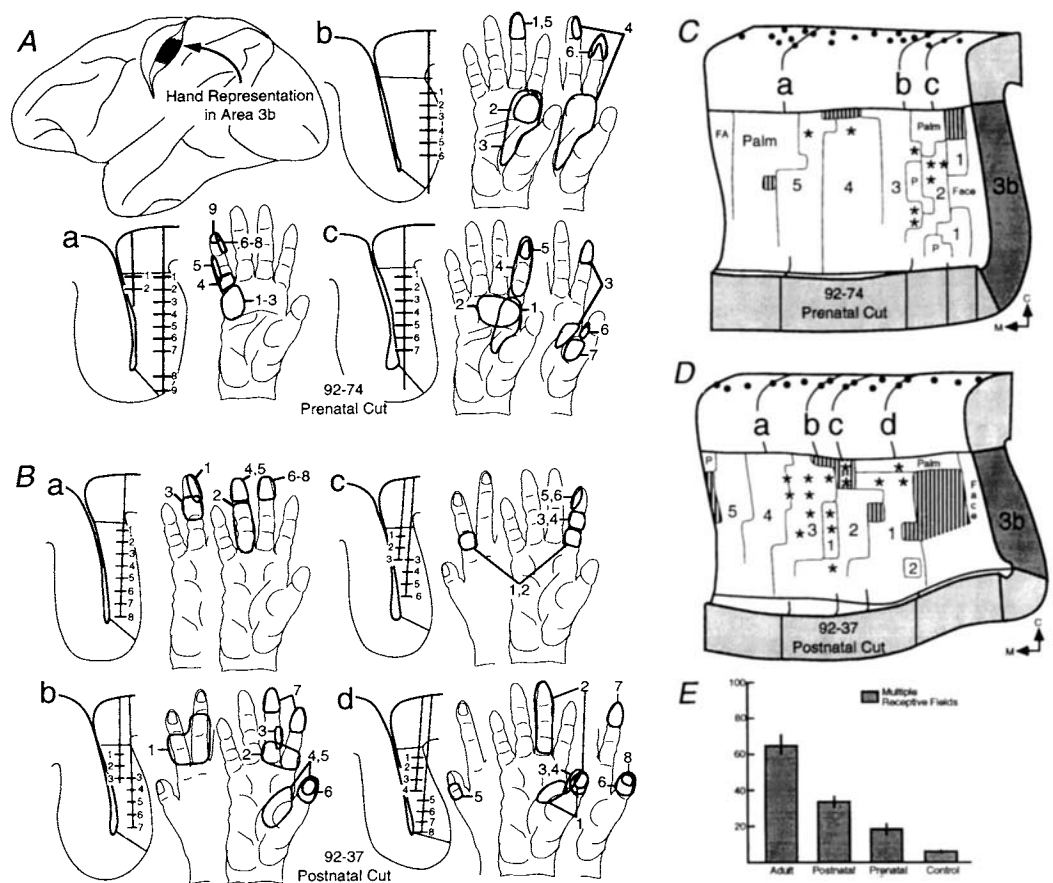
At some recording sites abnormal, 'multiple' receptive fields were identified. Multiple receptive fields are defined by equally robust responses from stimulation of two or more discrete skin regions. These were found in all monkeys regardless of the age at which the nerve was transected (indicated by asterisks in Figs 2 and 3); however, monkeys with nerve cut during prenatal development had significantly fewer recording sites with multiple receptive fields than monkeys with postnatal nerve cut or monkeys with nerve cut as adults (Fig. 2E).

In the monkeys that had the nerve cut and repair during fetal development, neurons occasionally presented other abnormal receptive field properties. One unusual feature was large, sub-optimal receptive field 'fringes'; such fringes are not found in normal monkeys⁷ and were not present in the monkeys with nerve cut during postnatal development or as adults⁴. These large fringes may result from excitatory activation of the widespread afferent inputs that has not been completely suppressed. The

second abnormal response property, which was occasionally seen after nerve cut in fetal monkeys, was a much reduced response to stimulation with the small diameter von Frey filaments, even when the filaments caused noticeable skin depression. Instead, larger-diameter stimuli (such as the diameter of the tip of a paper clip) produced robust responses in these monkeys even using stimulus pressures that caused no noticeable skin depression. This may indicate that summation of inputs from a larger skin area is necessary to drive the neurons, possibly as a result of a reduction in innervation densities in animals with prenatal nerve cut. In monkeys with postnatal nerve cut, as after nerve cut and repair in adults⁴, neurons responded robustly to fine-calibre filaments.

Summary maps of the hand in area 3b demonstrate the orderly representations of the digits innervated by the median nerve (Figs 2, 3). As in normal monkeys (Fig. 3a), the representation of each digit occupied a large, nearly continuous block of cortex, with the digit 1 representation most lateral and the other digits represented progressively more medially. There was some fracturing of the digit representations so that a given digit might be represented at more than one location in area 3b (Figs 2, 3). However, the

FIG. 2 Receptive fields and topographic organization in hand representations of area 3b following median nerve cut and repair in immature macaque monkeys. The schematic of the lateral view of the macaque brain (upper left) shows the approximate location of the hand representation (shaded area) in area 3b on the posterior bank of the central sulcus; the sulcus is shown partially opened for purposes of illustration. A, Receptive field progressions in area 3b of a monkey that had nerve cut and repair at 110 days of gestation. These receptive fields reflect the region which elicits the best response and does not include the suboptimal fringes. Outlines drawn on the hand figurines show the receptive field(s) at each numbered recording site in electrode tracks through area 3b (a, b and c). B, Receptive field progressions in area 3b of a monkey that had nerve cut and repair on postnatal day 10. C, D, Reconstructed maps of the hand representations in area 3b from the same animals as in A and B, respectively. Lower-case letters indicate the locations of the sections shown in the illustrations of the receptive field progressions (A and B). The thin lines in area 3b indicate the borders between the representations of the palm and each of the digits (indicated by numbers). Asterisks indicate locations of multiple receptive fields. Hatching indicates the representations of the dorsal skin of the hand. F, face; FA, forearm; P, palm. E, Bar graph of the average percent ($\pm$ s.e.m.) of multiple receptive fields after median nerve cut and repair at different developmental stages. METHODS. Low-impedance tungsten microelectrodes (1–2 megohms at 1 kHz) were used to record multiunit activity in the posterior bank of the central sulcus in monkeys anaesthetized with a combination of ketamine hydrochloride (20 mg kg⁻¹, i.m.) and xylazine hydrochloride (2 mg kg⁻¹, i.m.). Electrode penetrations were separated mediolaterally by 300–1,000 microns, depending on the surface vascular pattern, and within a



penetration recordings were made every 500 microns. At most of the recording sites, the sizes and locations of receptive fields were determined using calibrated von Frey filaments. The mapping procedures have been described in detail elsewhere^{4,7}. The rotated views of the posterior bank of the central sulcus were produced using a computerized image analysis system (R&M Biometrics). Calculations of multiple receptive fields were made both within the representations of the regenerated, median nerve (D1–D3) and within the normal portions of the hand representations (D4 and D5) (used as controls). Chi-square tests indicate that the differences in the average proportion of multiple receptive fields between the control and experimental values for each age group are significant ($P \leq 0.05$). Also, the differences between age groups are significant ($P \leq 0.05$). Data for the adult nerve cut animals are from the three cases described in ref. 4.

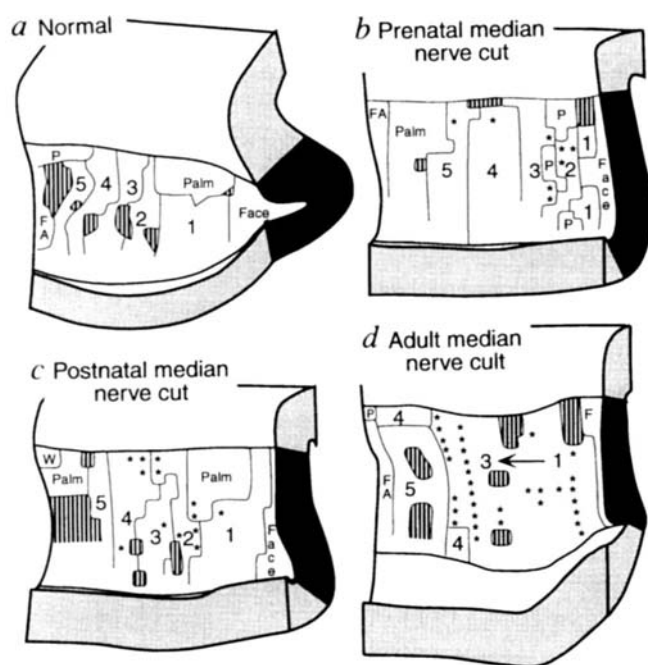


FIG. 3 Comparisons of the age-related effects of median nerve cut and repair on the organization of the hand representation in area 3b. *a*, Hand representation in a normal, adult macaque monkey. The map is redrawn from ref. 7. *b*, Hand representation after median nerve cut and repair on about 115 days gestation (*b*), at 10 days after birth (*c*), and in an adult macaque monkey (*d*); redrawn from ref. 4. Asterisks indicate the locations of multiple receptive fields. For further details, see Fig. 2 legend.

representations of the digits were situated in the appropriate lateromedial locations, so that the orderly progression from digit 1–5 was maintained. These progressions were consistent in all cases. In contrast, after nerve cut and repair in adult monkeys (Fig. 3*d*), neurons with receptive fields on digits 1, 2 and 3 were highly intermingled, so that there was only a crude topographic progression with digit 1 receptive fields concentrated more heavily in the lateral portion of the map and those from digits 2 and 3 concentrated progressively more medially.

These findings indicate that an orderly topographic representation of the reinnervated skin of the hand is established in primary somatosensory cortex, despite the profound topographic disorder in the sensory nerve termination patterns after early median nerve cut and repair. This contrasts with nerve cut and repair in adults, which results in seemingly permanent errors both in the peripheral nerve inputs and in the cortical representations of the reinnervated skin⁴. Thus, there are compensatory brain mechanisms in the somatosensory system during early development that correct for regeneration errors in the primary afferents. But what is the basis of the correction?

The creation of somatotopic order in cortex from disorder in the afferent inputs may depend on the selective pruning of developmentally exuberant cortical or even subcortical connections. Afferents from adjacent locations on the hand would be stimulated together, even in fetal life, producing correlated patterns of synaptic activity within subsets of the immature, presumably widespread⁸, connections. Connections would be preserved or lost based on correlations and disconnections in evoked activity^{9–11}, producing profound changes much like those shown in the visual system^{12–14} and in the rodent somatosensory system^{15–18} after early sensory manipulations.

Remarkably, the recovery of cortical somatotopy early in development appears to be largely independent of the order of the peripheral inputs. Instead, the pattern of sensory stimulation appears to be the most essential cue to establish appropriate

neighbour relationships in the brain. This implies that the critical factor in the treatment of sensory nerve transection in children is not the accuracy of nerve realignment, as has long been assumed¹⁹, but rather is sensory rehabilitation. □

Received 19 October 1995; accepted 9 February 1996.

1. Buchthal, F. & Kuhl, V. J. *Neurol. Neurosurg. Psych.* **42**, 436–451 (1979).
2. Dellon, A. L. *Evaluation and Sensibility and Re-education of Sensation in the Hand* (Williams and Wilkins, Baltimore, 1981).
3. Hallin, R. G., Wiesenfeld, Z. & Lindblom, U. *Exp. Neurol.* **73**, 90–106 (1981).
4. Florence, S. L., Garrahy, P. E., Wall, J. T. & Kaas, J. H. *Cerebral Cortex* **4**, 391–407 (1994).
5. Fitzgerald, M. J. *comp. Neurol.* **240**, 407–413 (1985).
6. Jain, N., Florence, S. L. & Kaas, J. H. *Dev. Brain Res.* **90**, 24–34 (1995).
7. Pons, T. P., Wall, J. T., Garrahy, P. E., Cusick, C. G. & Kaas, J. H. *Somatosens. Res.* **4**, 309–331 (1987).
8. Darian-Smith, I., Darian-Smith, I. & Cheema, S. S. *J. comp. Neurol.* **299**, 47–63 (1990).
9. Hebb, D. O. *Organization of Behavior* (Wiley, New York, 1949).
10. Stent, G. S. *Proc. natn. Acad. Sci. U.S.A.* **70**, 997–1001 (1973).
11. Changeux, J. P. & Danchin, A. *Nature* **264**, 705–712 (1976).
12. Hubel, D. H. & Wiesel, T. N. *Proc. R. Soc. Lond.* **198**, 1–59 (1977).
13. Shatz, C. J. & Sreter, D. W. A. *Rev. Neurosci.* **9**, 171–207 (1986).
14. Cline, H. T. & Constantine-Paton, M. in *Vision and Visual Dysfunction*, 11. *Development and Plasticity of the Visual System* (ed. Cronly Dillon, J. R.) (CRC, Boca Raton, 1991).
15. Kaas, J. H., Merzenich, M. M. & Killackey, H. P. A. *Rev. Neurosci.* **6**, 325–356 (1983).
16. Woolsey, T. A. in *Development of Sensory Systems in Mammals* (ed. Coleman, J.) (Wiley, New York, 1990).
17. O'Leary, D. D. M., Ruff, N. L. & Dyck, R. L. *Curr. Opin. Neurobiol.* **4**, 535–544 (1994).
18. Killackey, H. P., Rhoades, R. W. & Bennett-Clarke, C. A. *Trends Neurosci.* **18**, 402–407 (1995).
19. Bunnell, S. *Surgery of the Hand* (Lippincott, Philadelphia, 1956).

ACKNOWLEDGEMENTS. We thank J. Schall for comments on the manuscript, and L. Trice, J. Ives and J. Hudson for technical assistance. Supported by the NIH and by the J. F. Kennedy Center for Research on Human Development.

CORRESPONDENCE and requests for materials should be addressed to S.L.F. (e-mail address: florensi@ctrxv.vanderbilt.edu).

Long-term potentiation and functional synapse induction in developing hippocampus

Guylaine M. Durand, Yury Kovalchuk & Arthur Konnerth

Physiologisches Institut der Universität des Saarlandes, D-66421 Homburg, Germany

LONG-TERM potentiation (LTP) is a cellular mechanism that potentially underlies learning and memory¹. To test the hypothesis that LTP is involved in activity-dependent synapse formation, we combined whole-cell recordings and confocal microscopy to investigate hippocampal glutamatergic synapses at their earliest stages of development. Here we report that, during the first postnatal week, the hippocampal glutamatergic network becomes gradually functional owing to the transformation of precursor, pure NMDA (*N*-methyl-D-aspartate)-receptor-based synaptic contacts into conducting AMPA (α -amino-3-hydroxy-5-methylisoxazole-4-propionate)/NMDA-receptor-type synapses. This functional synapse induction is caused by an associative form of LTP, so it is input-specific and easily triggered experimentally by pairing presynaptic stimulation with postsynaptic depolarization. Our results challenge previous views that LTP occurs in the hippocampus only at later stages of development^{2–6} and that its induction requires dendritic spines⁷. They also provide direct evidence that LTP is important for the activity-dependent formation of conducting glutamatergic synapses in the developing mammalian brain.

Synapse formation in the brain involves several sequential processes, including the simultaneous growth of both the afferent axon and the corresponding target neuron into opposition, followed by the elaboration of structural specializations at the

contact region. Morphological studies have demonstrated that CA1 pyramidal neurons of the rat hippocampus are generated at embryonic day 17–19 (ref. 8). The growth of their apical dendrites leads to the rapid expansion of the dendritic layer, the so-called stratum radiatum, during a period lasting from embryonic day 22 (E22, the last intrauterine day in rats) to postnatal day 1 (P1).

We used whole-cell voltage-clamp recordings⁹ applied to CA1 pyramidal neurons in rat hippocampal slices to investigate the function of glutamatergic synapses at their earliest stages of development. Up to P2, stimulation of afferent fibres in the stratum radiatum produced virtually no synaptic responses when the cells were held around their resting potential (−60 mV). Surprisingly, however, synaptic responses could be detected in neurons that were held at depolarized membrane voltages, for example at +40 mV. On repolarization to −60 mV, these synapse became 'silent' (Fig. 1a, top). Their voltage dependence and their sensitivity to D-2-amino-5-phosphovalerate (AP5) identified the excitatory postsynaptic currents (e.p.s.c.s) recorded at +40 mV as NMDA receptor-mediated e.p.s.c.s. These and other similar experiments ($n = 9$) lead us to conclude that at the earliest developmental stages glutamatergic synapses consist of pure NMDA-mediated synaptic contacts, as indicated in the schematic diagram representing the CA1 pyramidal cell layer (Fig. 1a). In contrast, at P6, synaptic transmission was frequently detected at the resting potential, as in the adult¹⁰, owing to synapses containing both non-NMDA and NMDA receptors (Fig. 1b). We use the term conducting synapses to refer to these synapses that are functional at resting potential.

An analysis of synaptic responses throughout the first postnatal week indicates that the stage of predominant NMDA-mediated synaptic contacts is short lived (Fig. 1c). At birth (P0), no synaptic inputs were detected, whereas at P1 and P2 most (82%) of the pure NMDA receptor-mediated e.p.s.c.s were found. The percentage of pure NMDA synapses over that of total synapses

decreased during the first postnatal week (Fig. 1c). Remarkably, pure NMDA synapses have also been reported to be present during the second and third postnatal week^{11,12}.

The immature glutamatergic synapses found early in development were not only functionally but also morphologically different from the mature ones (Fig. 2). In recordings combining whole-cell patch clamp and confocal microscopy, we identified the site of synaptic contact by recording the e.p.s.c. and the corresponding Ca^{2+} signal simultaneously (Fig. 2b, c). A detailed morphological examination of this and six other neurons analysed using the same approach demonstrated the absence of any spine-like structure at the site of Ca^{2+} signalling. In a series of similar experiments performed on slices from rats that were older than P12, Ca^{2+} signalling associated with synaptic stimulation was always most prominent in spines (not shown). These results are also consistent with two-photon imaging results¹³ obtained at later stages of development. Moreover, detailed electron microscopical studies showed that, in CA1, the few synapses present at P1 were located directly on dendritic shafts or on stubby protrusions and were proposed to be immature spiny synapses¹⁴.

To test whether pure NMDA-mediated synaptic contacts are precursors of the mature conducting glutamatergic synapses, we identified 'silent' inputs by stimulating afferent fibres, either at +40 mV in standard (1 mM Mg^{2+}) or at −70 mV in Mg^{2+} -free perfusion solution (Fig. 3a, b). At negative membrane potentials in standard solution, no synaptic responses were detected. For both experimental protocols (Fig. 3a, b), pairing of 20 afferent stimuli (at 0.2 Hz) with postsynaptic depolarization to 0 mV reliably induced synaptic responses at previously silent synaptic contacts. We termed this process of rapid, activity-dependent development of conducting synapses 'functional synapse induction'.

Such functional synapse induction occurred in most cases less than a minute after pairing ($n = 15$). In six attempts, pairing failed

FIG. 1 High incidence of pure NMDA-receptor-mediated synaptic transmission in newborn rats. **a**, Whole-cell voltage-clamp recording from a CA1 pyramidal neuron in a hippocampal slice from a 2-day-old-rat (P2). Afferent single-shock stimulation evoked e.p.s.c.s detected at a holding potential of +40 mV, but not at −60 mV (control). These e.p.s.c.s were completely and reversibly blocked by 50 μM AP5. Bottom, the experimental result illustrated schematically. **b**, e.p.s.c.s recorded in a hippocampal CA1 pyramidal neuron from a 6-day-old-rat. In addition to the NMDA receptor component, these e.p.s.c.s also reveal an AMPA component at −60 mV (control) that is reversibly blocked by 5 μM CNQX. Bottom, a model of the corresponding synapse with colocalized AMPA and NMDA receptors. Recordings in **a** and **b** display superpositions of 5 consecutive synaptic responses. Arrowheads indicate the time of stimulation. **c**, The proportion of pure NMDA synapses decreases during the first week of postnatal development. Bars represent the percentage of synapses producing only NMDA receptor-mediated e.p.s.c.s over the total number of synapses (NMDA-mediated + conducting synapses). The numbers in parentheses above the bars indicate the total number of synaptic inputs tested per age. At P0, 400 stimulation sites in 8 cells were tested as described below. Even with the minimal stimulation used here, there is no certainty of stimulating a single input axon.

METHODS. Hippocampal slices (300 μm thick) from Wistar rats (P0–P12) were prepared as previously described^{9,23}. The slices were perfused with standard solution containing (in mM): 125 NaCl, 2.5 KCl, 1.25 NaH_2PO_4 , 26 NaHCO_3 , 2 CaCl_2 , 1 MgCl_2 , 20 glucose, bubbled with 95% O_2 , 5% CO_2 . The pipette solution contained (in mM): 135 CsCl, 0.6 EGTA, 20 TEA, 2 MgCl_2 , 10 HEPES, 0.4 Na-GTP, 4 $\text{Na}_2\text{-ATP}$, pH 7.3. Combined whole-cell recordings and confocal Ca^{2+} imaging were performed as described^{24,25}. For synaptic stimulation of afferent fibres, voltage pulses (200 μs duration, 20 V) delivered at 0.1 Hz through a pipette placed in the stratum radiatum. For the screening experiments performed at P0, 3 consecutive stimuli (20 Hz) were given every 10 s in 50 different locations per neuron to increase the probability of finding inputs. In all experiments, GABA_A receptor-mediated transmission was blocked by bicuculline (10 μM) added to the standard solution. APV, CNQX and cyclothiazide were also added to the standard solution when needed. Recordings were made at room temperature (22–24 °C).

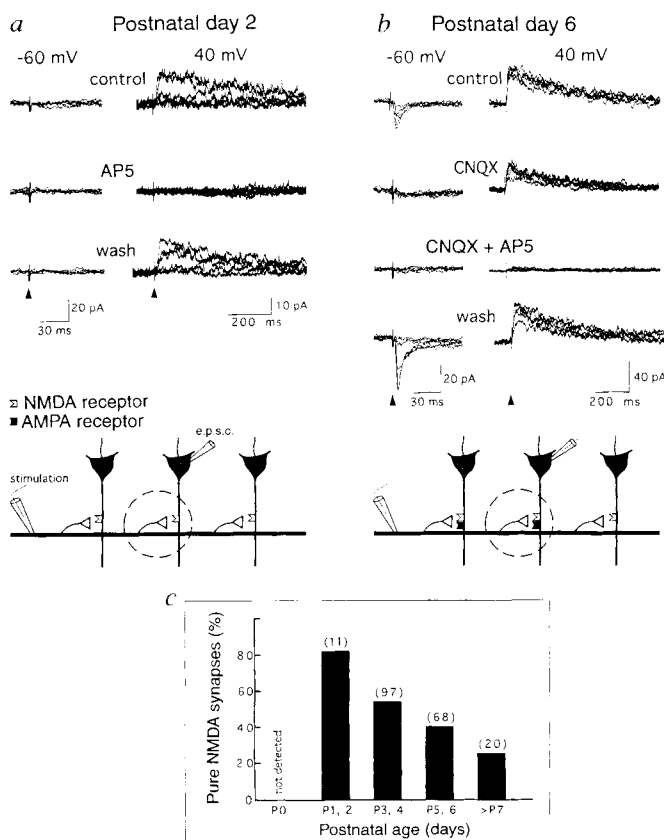
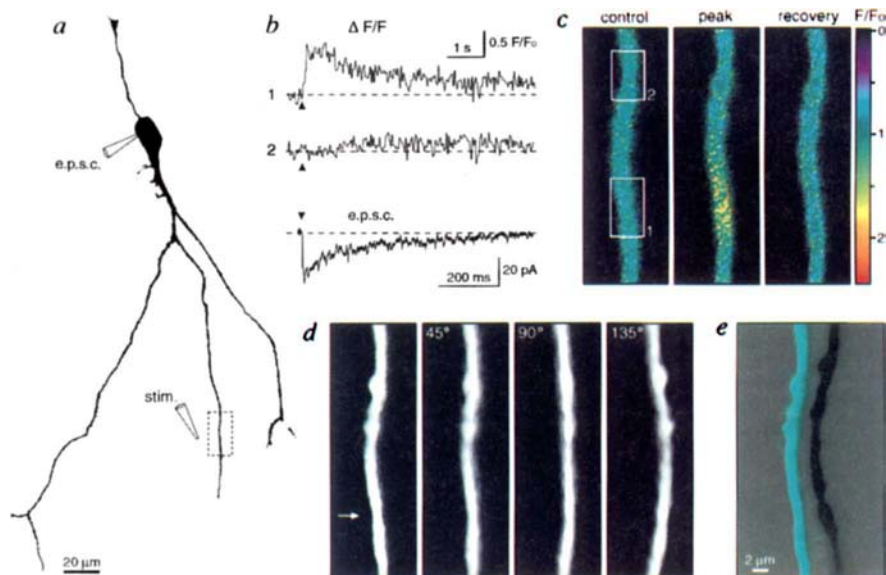
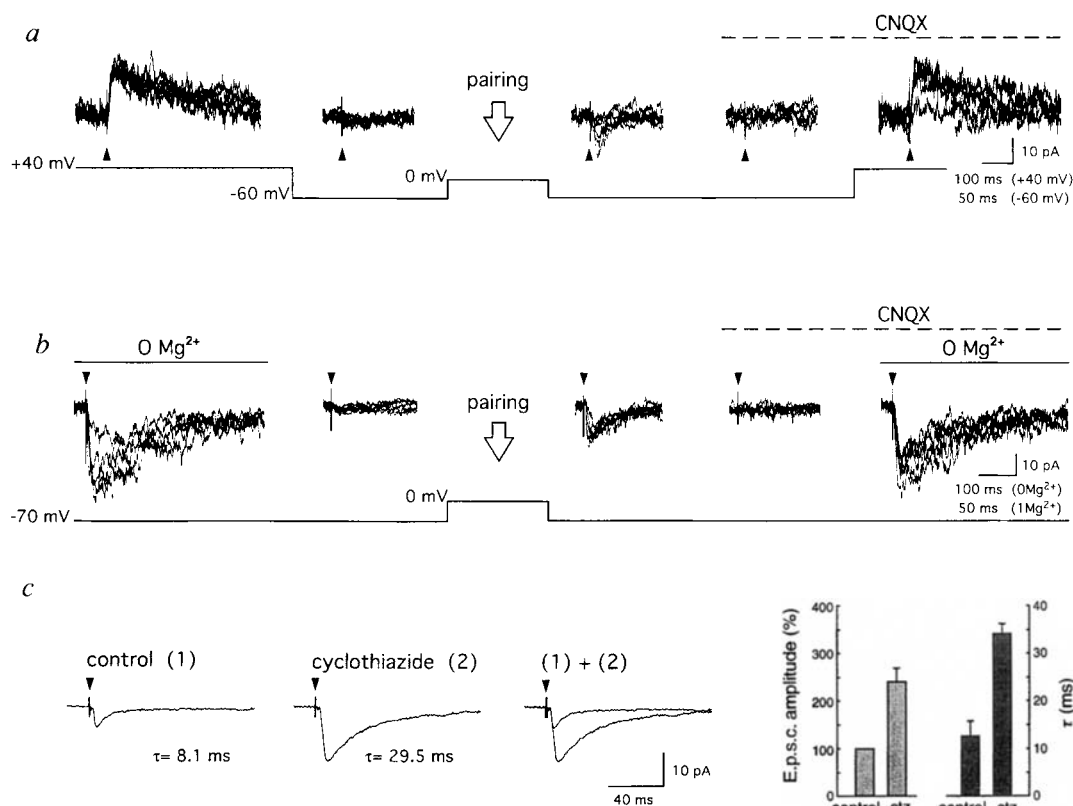


FIG. 2 Evidence for non-spiny glutamatergic synapses in CA1 pyramidal neurons in newborn rats. *a*, CA1 pyramidal neuron (P2) reconstructed from serial confocal sections of a cell injected with Ca Green 1 (250 μ M) through the patch pipette. The region in the dotted box represents the dendritic area depicted in *c*–*e*. The schematic pipettes indicate the sites of whole-cell recording (e.p.s.c.) and synaptic stimulation (stim.), respectively. *b*, Combined whole-cell recording of an e.p.s.c. (bottom) and corresponding Ca^{2+} -dependent fluorescence signals (traces 1 and 2) obtained by confocal Ca^{2+} imaging. The synaptic response was recorded in Mg^{2+} -free standard solution (at a holding voltage of -60 mV) and represents an NMDA receptor-mediated e.p.s.c. Perfusion with the standard saline (1 mM Mg^{2+}) abolished both the e.p.s.c. and the Ca^{2+} transients. *c*, Pseudocolour images representing the changes in Ca^{2+} -sensitive fluorescence at selected time points during the synaptic response shown in *b*. Each image is an average of 8 consecutive images. *d*, Results obtained from the computer-based reconstruction and restoration of the dendritic segment from *c*, viewed from 4 different angles of rotation (0° , 45° , 90° and 135° , respectively). Note the absence of any spine-like protrusion near the arrow, which indicates the presumed site of synaptic contact as deduced from the peak Ca^{2+} signal (box 1 in *c*). *e*, Three-dimensional representation of the same dendritic segment. Reconstructions and restorations in *d* and *e* were made from 60 serial optical sections



(step size 270 nm) using Imaris software (Bitplane Zürich, Switzerland). The point-spread function for the Tikhonov–Miller algorithm was generated using fluorescent beads of diameter $\sim 0.1 \mu\text{m}$ (Molecular Probes, Eugene, USA).

FIG. 3 Activity-dependent synapse induction in the developing hippocampus. *a*, *b*, Two approaches used to identify 'silent' synaptic inputs. *a*, First, a synaptic input mediating NMDA e.p.s.c.s was identified by holding the potential at $+40$ mV. Stimulation of the same input at -60 mV confirmed that the input was 'silent'. At the time point marked with an open arrow, the pairing of stimulation of afferent pre-synaptic fibres (20 stimuli every 5 s) with postsynaptic depolarization (to 0 mV for 2 min) 'induced' synaptic transmission at -60 mV. The newly induced e.p.s.c. component was blocked by CNQX (5 μM). The last group of traces in this row is similar to that registered at the beginning, indicating that NMDA-receptor based transmission was not affected by synapse induction. *b*, The second approach involved identifying 'silent' NMDA-synapses in Mg^{2+} -free perfusion solution. The protocol used in *a* was then applied to produce synapse induction. Each recording in *a* and *b* represents a superposition of 6 consecutive responses obtained at a 10-s interval. *c*, E.p.s.c.s recorded at -60 mV before (1) and after (2) addition of 100 μM cyclothiazide. The increase in amplitude and the prolongation of



the decay time constant (τ) indicate that the e.p.s.c.s were produced by activation of AMPA receptors. Each trace is an average of 20 consecutive sweeps. The bar graph summarizes results from 4 experiments; error bars represent s.e.m.

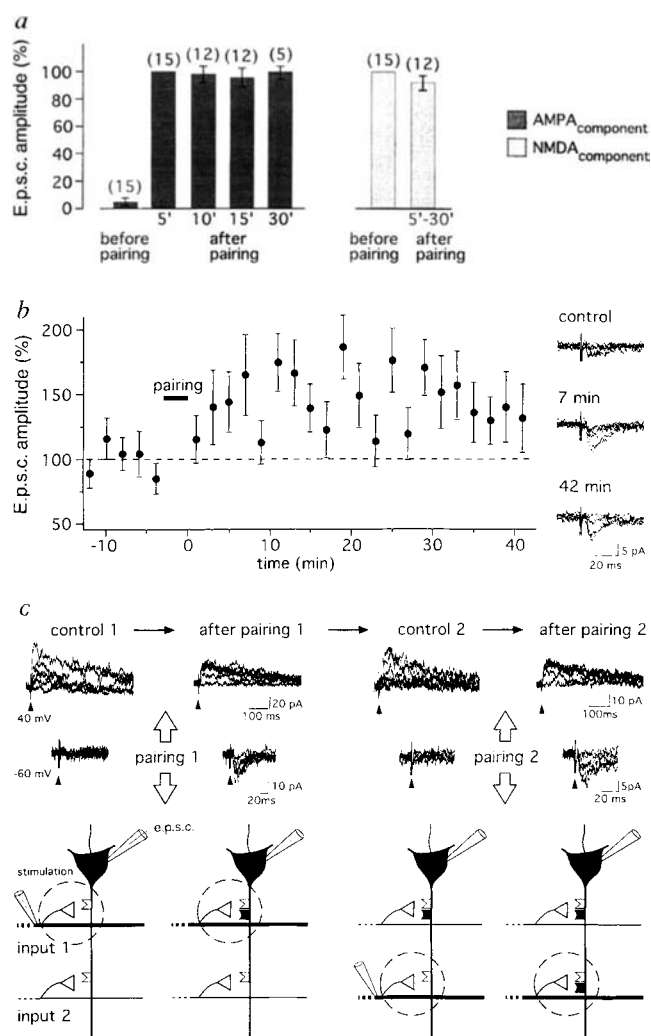


FIG. 4 Persistence of the efficacy of newly 'induced' synapses and their input specificity. **a**, Amplitudes of e.p.s.c.s before and at various intervals after pairing. The bar graph summarizes the results of 15 experiments, of which 11 represent cases of synapse induction at 'silent' inputs and 4 of LTP at conducting synapses in rats 3–5 days old. Note the absence of any potentiation of the synaptic NMDA component (right). The apparent reduction of the NMDA component after pairing was not significant ($P < 0.05$, $n = 15$, t -test). The number of neurons monitored for each time window is indicated on top of bars. For the AMPA components, the e.p.s.c. amplitudes were normalized to the averaged first responses ($n = 15$) recorded after pairing. The NMDA components were normalized to the averaged NMDA response ($n = 15$) registered as control before pairing. **b**, LTP of a synaptic input that was conducting at resting potential at P2. The same pairing protocol was used as for synapse induction. Each data point represents the mean amplitude of 12 consecutive e.p.s.c.s evoked at 0.1 Hz. Right, superpositions of 6 consecutive traces recorded at various times during the experiment. Error bars in **a** and **b** represent s.e.m. **c**, Sequential e.p.s.c. recordings demonstrating synapse induction at two distinct inputs of the same neuron (P3) by using the approach illustrated in Fig. 3a. After 'inducing' synaptic transmission at the first input, the stimulation pipette was placed at a new site (about 100 μ m away from the first one) identifying another 'silent' input (control 2). Pairing depolarization and afferent stimulation of the second input (pairing 2) 'induced' an additional conducting synapse. Bottom, schematic illustration of the temporal progression of synapse induction at the two different inputs. The areas in broken circles represent the corresponding sites involved in synapse induction. Similar results were obtained from 6 cells.

to induce conducting synapses. The newly 'induced' e.p.s.c.s were completely blocked by 6-cyano-7-nitroquinoxaline-2,3-dione (CNQX), identifying them as non-NMDA receptor-mediated e.p.s.c.s (Fig. 3a, b). To assess further the pharmacological identity of the newly formed non-NMDA component, we used cyclothiazide, which removes desensitization of AMPA receptors but has no effect on kainate receptors^{15,16}. In the presence of cyclothiazide, the e.p.s.c. component recorded at -60 mV was potentiated by $241 \pm 28\%$ ($n = 4$, mean $\pm$ s.e.m.) and its decay time constant was more than three times longer than the control (Fig. 3c). This indicates that the earliest glutamatergic synaptic responses detected at resting potential are predominantly mediated by AMPA receptors.

Functional synapse induction and LTP are mechanistically similar processes (Fig. 4). The amplitude of the 'de novo' induced AMPA receptor-mediated e.p.s.c.s was constant throughout the remaining time of whole-cell recording (up to 30 min; Fig. 4a). This suggests that, once established, conducting synapses become a persistent element of the neuronal circuit. The NMDA receptor-mediated component seemed not to be affected by synapse induction. Its amplitude was nearly the same before and after pairing (Fig. 4a), suggesting that synapse induction is a postsynaptic process. As for LTP^{17,18}, synapse induction was blocked when buffering intracellular Ca^{2+} by 10 mM EGTA ($n = 4$). Moreover, functional synapse induction could not be produced when NMDA receptors were blocked by perfusing the slice with 50 μ M AP5 ($n = 3$). Further, the activity-dependent induction of the synaptic AMPA response was input-specific (Fig. 4c). Stimulation of a second 'silent' input to the same cell leads to the

induction of another synapse, in response to the standard pairing protocol applied to the second input. This finding demonstrates that postsynaptic depolarization alone is not sufficient for synapse induction (see also Fig. 3a, b). Thus functional synapse induction, like LTP in adults¹⁹, requires the coincident activity of the pre- and the postsynaptic site, in line with Hebb's rule of association²⁰.

Our results contrast with those obtained in other studies which suggested that LTP is not present before P5 (refs 3–6) and that the only form of synaptic plasticity detectable during the first postnatal week (P3–7) is long-term depression². Under our recording conditions, synapse induction and LTP (of the occasionally found conducting synapses) could be induced as early as P2 by pairing postsynaptic depolarization with afferent stimulation (Fig. 4b). This finding is consistent with reports suggesting that a basic mechanism of LTP in older (10–12 days old) rats is based on a transformation of silent synapses into functional ones owing to an increase in their AMPA-type responsiveness^{11,12}.

In conclusion, these results show that, at their earliest stages of development, hippocampal glutamatergic synapses involve almost exclusively NMDA receptors, forming a neuronal network that is silent at resting potential. The NMDA-receptor based network, which seems to form independently of any sensory activity, may act as a blueprint for the developing functional hippocampal circuit. At the preformed 'silent' contacts, electrical activity can 'induce' functionally stable synapses that consist postsynaptically of colocalized AMPA and NMDA receptors^{21,22}. Thus our findings provide direct evidence for the involvement of LTP, or a closely related process, in the activity-dependent formation of conducting glutamatergic synapses in the mammalian hippocampus. □

Received 23 November 1995; accepted 14 March 1996.

- Bliss, T. V. P. & Collingridge, G. L. *Nature* **361**, 31–39 (1993).
- Bolshakov, V. Y. & Siegelbaum, S. A. *Science* **264**, 1148–1152 (1994).
- Dudek, S. M. & Bear, M. F. *J. Neurosci.* **13**(7), 2910–2918 (1993).
- Bekenstein, J. W. & Lothman, E. W. *Dev. Brain Res.* **63**, 245–251 (1991).
- Harris, K. M. & Teyler, T. J. *J. Physiol., Lond.* **346**, 27–48 (1984).
- Baudry, M., Arst, D., Oliver, M. & Lynch, G. *Brain Res.* **227**(1), 37–48 (1981).
- Harris, K. M. & Kater, S. B. *A. Rev. Neurosci.* **17**, 341–371 (1994).
- Bayer, S. A. *J. comp. Neurol.* **190**, 87–114 (1980).
- Edwards, F., Konnerth, A., Sakmann, B. & Takahashi, T. *Pflügers Arch.* **414**, 600–612 (1989).
- Hestrin, S., Nicoll, R. A., Perkel, D. J. & Sah, P. *J. Physiol., Lond.* **422**, 203–225 (1990).
- Liao, D., Hessler, N. A. & Malinow, R. *Nature* **375**, 400–404 (1995).
- Isaac, J. T. R., Nicoll, R. A. & Malenka, R. C. *Neuron* **15**, 427–434 (1995).

13. Yuste, R. & Denk, W. *Nature* **375**, 682–684 (1995).
14. Steward, O. & Falk, P. M. *J. comp. Neurol.* **314**, 545–557 (1991).
15. Partin, K. M. et al. *Neuron* **11**, 1069–1082 (1993).
16. Wong, L. A. & Mayer, M. L. *Molec. Pharmacol.* **44**, 504–510 (1993).
17. Lynch, G. et al. *Nature* **305**, 719–721 (1983).
18. Malenka, R. C., Kauer, J. A., Zucker, R. S. & Nicoll, R. S. *Science* **242**, 81–84 (1988).
19. Gustafsson, B. & Wigström, H. *Trends Neurosci.* **11**, 156–162 (1988).
20. Hebb, D. O. *The Organization of Behaviour* (Wiley, New York, 1949).
21. Bekkers, J. M. & Stevens, C. F. *Nature* **341**, 230–233 (1989).
22. Konnerth, A., Keller, B. U. & Lev-Tov, A. *Pflügers Arch* **417**, 285–290 (1990).
23. Konnerth, A. *Trends Neurosci.* **13**, 321–323 (1990).
24. Eilers, J., Augustine, G. J. & Konnerth, A. *Nature* **373**, 155–158 (1995).
25. Eilers, J., Callewaert, G., Armstrong, C. & Konnerth, A. *Proc. natn. Acad. Sci. U.S.A.* **92**, 10272–10276 (1995).

ACKNOWLEDGEMENTS. We thank E. Hanse, O. Garaschuk and T. Plant for comments on the manuscript, and N. Rothgerber and E. Eilers for technical assistance. The work was supported by grants from the BMBF and the DFG (A.K.). G.M.D. was supported by the Alexander von Humboldt-Stiftung and the Human Frontier Science Program.

CORRESPONDENCE and requests for materials should be addressed to A.K. (e-mail: konnerth@rz.uni-sb.de).

Critical protective role of mast cells in a model of acute septic peritonitis

Bernd Echtenacher, Daniela N. Männel
& Lothar Hültner*

Institut für Pathologie/Tumorimmunologie, Universität Regensburg,
Franz-Josef-Strauss-Allee, D-93042 Regensburg, Germany

* GSF-Institut für Experimentelle Hämatologie, Marchioninistrasse 25,
D-81377 Munich, Germany

MAST cells play a detrimental role in IgE-dependent allergic reactions. In contrast, a protective function for mast cells has been proposed on the basis of some worm infection models. No reports exist on the *in vivo* significance of these cells in bacterial infections^{1,2}. Here we use congenitally mast-cell-deficient *W/W^v* mice and normal *+/+* littermates^{3,4} to analyse the role of mast cells in a model of acute septic peritonitis (caecum ligation and puncture (CLP)). Following CLP, *W/W^v* mice showed a significantly increased mortality compared to *+/+* mice. The selective reconstitution of *W/W^v* mice with cultured *+/+* mast cells substantially protected them from the lethal effects of CLP, whereas an anti-tumour-necrosis-factor (TNF) antibody injected immediately after CLP completely suppressed this protection. Our results reveal a previously unrecognized protective role of mast cells and mast-cell-derived TNF in acute bacterial peritonitis.

Mast-cell-deficient (*WB* × *C57Bl/6*) *F₁*-*W/W^v* mice were significantly more sensitive to an experimentally induced acute bacterial peritonitis than their normal congenic (*WB* × *C57Bl/6*) *F₁*-*+/+* littermates (hereafter called *W/W^v* and *+/+*, respectively). This resulted in 100% mortality of *W/W^v* mice within 5 days following caecal ligation and puncture (CLP), as opposed to only 25% mortality in *+/+* mice (Fig. 1). Finally, 8 of 12 CLP-treated *+/+* mice fully recovered from the symptoms of acute peritonitis and remained healthy during the observation period of 40 days. Elegant experiments by Kitamura and colleagues have demonstrated the adoptive transfer of cultured bone-marrow-derived mast cells (BMMC) generated from normal *+/+* mice: they were able selectively to reconstitute locally both mucosal-type and connective-tissue-type (or serosal) mast cells in *W/W^v* mice without correcting the macrocytic anaemia observed in these mutants (refs 5,6; reviewed in refs 3,4). In a pilot study, we transplanted intraperitoneally (i.p.) graded numbers of *+/+* BMMC (1.7×10^6 , 5×10^6

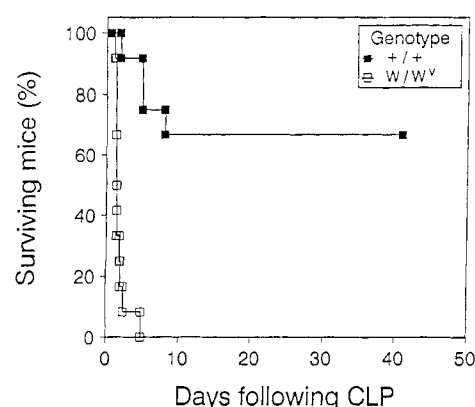


FIG. 1 Highly increased mortality of mast-cell-deficient *W/W^v* mice after CLP, compared to identically treated age- and sex-matched normal littermates (*+/+*) (12 mice per group) ($P < 0.0001$ for *+/+* versus *W/W^v*; log rank statistic). *W/W^v* and *+/+* mice were generated from heterozygous breeding pairs (*W/+* × *W/+*) originally obtained from The Jackson Laboratory (Bar Harbor, ME). The experimental surgical procedure, including anaesthesia of mice, has been described⁷. Briefly, in deeply narcotized mice, the distal end of the caecum (about 40% of the total caecal length) was ligated and one puncture was set with a 0.9 mm needle. Note that in control groups of *W/W^v* and *+/+* mice not subjected to CLP, the spontaneous mortality was zero during the observation period of 40 days. Results were similar in two other independent experiments.

or 15×10^6 cells per mouse; *in vitro* age: 5 weeks; > 99% Alcian blue positive) into *W/W^v* mice and subsequently performed a quantitative and qualitative analysis of mast cells recovered from the peritoneal cavities of *W/W^v* mice at different times after reconstitution. Although we used a different protocol for culturing BMMC, the results shown in Table 1 confirmed the selective reconstitution of *W/W^v* mice with mast cells through i.p. transplantation of cultured *+/+* BMMC (refs 3–6), and in addition revealed that 1.7×10^6 *+/+* BMMC transplanted per mouse 20 days before CLP were sufficient to give protection to most of the reconstituted *W/W^v* mice (data not shown). Under our culture conditions (using interleukins IL-3 and IL-4 as mast-cell growth and differentiation factors), only a few BMMC (< 10%) stained positively with safranin, a dye specific for mature serosal-type mast cells. However, 20 days after i.p. injection of these cultured BMMC into *W/W^v* recipients, 45–55% of the mast cells recovered from the peritoneal cavities of reconstituted *W/W^v* mice stained positively with this dye, indicating that a substantial proportion of these BMMC (or their progeny) had acquired a more mature serosal-like mast-cell phenotype.

TABLE 1 Erythrocyte counts, haematocrit and peritoneal mast cells in *+/+* mice, *W/W^v* mice and mast-cell-reconstituted *W/W^v* mice (*W/W^v* (BMMC))

Mice	Mean values ± s.d.		
	Erythrocytes μl^{-1} ($\times 10^6$)	Haematocrit (%)	Peritoneal mast cells per mouse ($\times 10^4$)
<i>+/+</i> ($n = 4$)	9.05 ± 1.21	41.25 ± 3.59	21.9 ± 4.3
<i>W/W^v</i> (BMMC) ($n = 4$)	4.94 ± 0.49	30.40 ± 2.88	31.4 ± 9.9
<i>W/W^v</i> ($n = 5$)	4.81 ± 0.44	29.75 ± 2.87	0.0 ± 0.0

Mast-cell-reconstituted *W/W^v* mice were analysed 21 days after i.p. transplantation of 5×10^6 *+/+* BMMC. Mast cells were recovered from peritoneal lavages. The differences in the values for the number of erythrocytes and haematocrit were statistically significant ($P < 0.001$) between *+/+* mice and the *W/W^v* or *W/W^v* (BMMC) group of mice, whereas there was no statistically significant difference of these values between *W/W^v* and *W/W^v* (BMMC) mice. 100% of the peritoneal mast cells recovered from *+/+* mice and 51.2 ± 5.9% of the peritoneal mast cells recovered from *W/W^v* (BMMC) mice stained positively with safranin.

The proinflammatory cytokine TNF had previously been identified in the CLP model of sepsis as a critical protective component of the host defence system which is able to manage the potentially lethal effects of intestinal bacteria invading the peritoneal cavity and resulting in bacteraemia⁷. We therefore investigated whether the protective effect exerted by transplanted BMDC or their progeny was also dependent on TNF. The results shown in Fig. 2 proved that i.p. transplantation of normal $+/+$ BMDC into W/W^v mice 20 days before CLP could substantially protect reconstituted W/W^v mice from the otherwise lethal effects of CLP. As a single injection of a blocking anti-TNF antibody immediately after CLP treatment of mast-cell-transplanted W/W^v mice completely suppressed this protection, we concluded that this critical protective role of mast cells in CLP-induced acute septic peritonitis was dependent on endogenous TNF (Fig. 2).

The mast cell is an important cellular source of different multifunctional cytokines⁸. Moreover, the mast cell is apparently the only resident cell type capable of storing TNF in cytoplasmic granules that is rapidly released (within 10 to 20 min) of challenge^{9,10}. We consider that this unique property of mast cells may be critically important here. Other resident peritoneal cells that possibly contribute to TNF secretion later on in the inflammatory reaction (macrophages for example) are expected to start with TNF production after a significant time delay, whereas TNF-producing neutrophils have to be recruited from the circulation. As the number of peritoneal macrophages does not differ between W/W^v mice and normal littermates (our unpublished results and ref. 11), this cell type is an unlikely candidate to explain the observed differences. We have not yet identified the mechan-

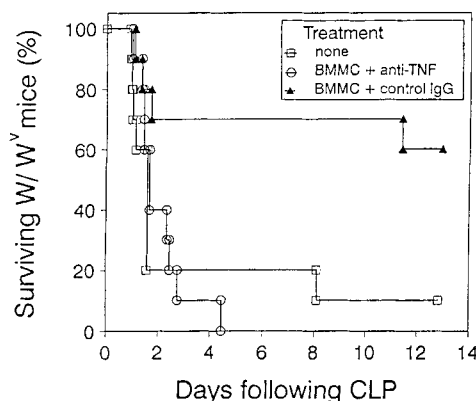


FIG. 2 Protection of W/W^v mice from the potentially lethal effects of CLP through transplantation of cultured $+/+$ BMDC and reversal of protection by anti-TNF. The CLP-induced mortality rates of different groups of age- and sex-matched W/W^v mice (10 mice per group) have been compared. 20 days before CLP, 20 W/W^v mice had been transplanted with normal cultured $+/+$ BMDC (5×10^6 cells injected i.p. per mouse). Ten of these mast-cell-reconstituted W/W^v mice were injected immediately following CLP with a saturating dose of 100 μ g per mouse i.p. of the neutralizing monoclonal rat anti-mouse TNF antibody V1q (ref. 7); another group of mast-cell-reconstituted W/W^v mice received a normal rat IgG (Sigma; 100 μ g per mouse i.p.). A third group of 10 CLP-treated control W/W^v mice (not reconstituted with $+/+$ BMDC) was injected with PBS (0.5 ml per mouse, i.p.) immediately after CLP ($P < 0.005$ for BMDC + anti-TNF versus BMDC + control IgG; $P < 0.002$ for none versus BMDC + control IgG). BMDC were generated essentially as described^{28,29}. In brief, normal $+/+$ bone marrow cells were initially grown in RPMI 1640 medium containing 20% FCS, 2 mM L-glutamine, 100 U ml⁻¹ penicillin-streptomycin, 10^{-5} M α -thioglycerol and 10 ng ml⁻¹ recombinant (r) murine (mu) IL-3. After 14 days, all non-adherent cells were transferred to fresh culture media additionally supplemented with rmuIL-4 (10 ng ml⁻¹ final). After two additional weekly refeedings (renewing both culture flasks and media), apparently pure populations of $+/+$ mast cells (*in vitro* age: 5 weeks; > 99% Alcian blue positive, including 7% safranin-positive cells) were obtained and used to reconstitute W/W^v mice. These results were confirmed in two further independent experiments.

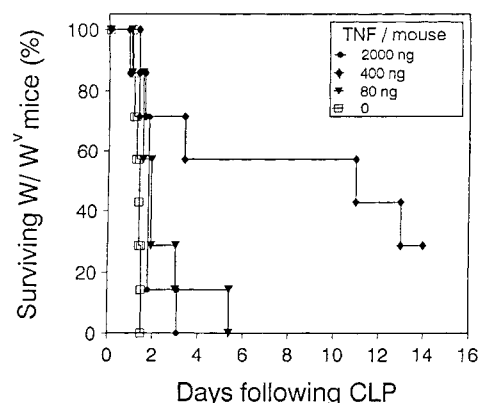


FIG. 3 A single injection of TNF immediately after CLP can exert a protective function in W/W^v mice not reconstituted with mast cells. Groups of age- and sex-matched W/W^v mice (7 mice per group) were treated with CLP immediately followed by i.p. injection of graded doses of rmuTNF (2,000, 400 or 80 ng per mouse) or control diluent (0.5 ml PBS per mouse) ($P < 0.002$ for 400 ng versus 0; $P < 0.03$ for 400 ng versus 2,000 ng; $P < 0.04$ for 400 ng versus 80 ng). Purified rmuTNF was provided by Knoll AG (Ludwigshafen, Germany). In dose-response studies with rmuTNF, injected i.p. into groups of normal mice at the time of initiation of CLP, we repeatedly confirmed the protective effects of low doses of TNF (200–600 ng per mouse), whereas doses of > 1,000 ng per mouse of TNF were ineffective or even detrimental (data not shown). Similar findings had been published in a CLP model of sepsis in rats³⁰. Comparable results were obtained in a second independent experiment.

ism(s) that leads to the release of preformed TNF from mast-cell stores in CLP-induced acute bacterial peritonitis, but in the light of recent findings our experiments will focus on the possible involvement of endotoxin¹² and/or substance P (refs 13–15) as known mast-cell activators.

Our present results do not exclude the possibility that mast cells may be directly involved in the phagocytosis of bacteria^{16–18}. Indeed, mast cells bind preferentially to type-1 fimbriae present on *Escherichia coli* and other species of enteric bacteria^{19,20}. Adherence of such bacteria to peritoneal mast cells can cause mediator release¹⁹ followed by phagocytosis and killing²⁰. But we favour an important role for mast cells in the rapid recruitment of neutrophils from the circulation and in the functional activation of these classical phagocytes by mechanisms involving TNF and possibly other mast-cell-derived mediators and cytokines^{11,21}. Interestingly, depletion of neutrophils significantly enhances CLP-induced mortality in normal mice (B.E., unpublished results), a finding that agrees with many other infection models. As shown in Fig. 3, i.p. injection of 400 ng recombinant murine (mu) TNF per mouse also had a significant protective effect in CLP-treated W/W^v mice not reconstituted with mast cells. A higher dose (2,000 ng per mouse) of recombinant muTNF injected i.p. directly after CLP treatment of W/W^v mice was not beneficial (Fig. 3). This failure may be readily explained by the well-known local and systemic toxicity of TNF in experimental animals²², which is even more pronounced in infected animals²³. Although the protective effect of 400 ng per mouse of recombinant muTNF injected i.p. in CLP-treated non-reconstituted W/W^v mice proved statistically significant (Fig. 3), the mast-cell-reconstitution protocol was superior with regard to the long-term survival of CLP-treated W/W^v mice (Fig. 2). This finding may be readily explained by the short half-life of injected TNF *in vivo* and by the possibility that besides TNF other mast-cell-derived mediators may synergistically support the phalanx of innate humoral and cellular defence components aimed at localizing bacterial infections. An important mechanism for this localization is fibrin formation²⁴, which was dependent on mast cells in another model system¹⁵. Interestingly, TNF depletion *in vivo* not only led to

significantly higher CLP-induced mortality rates in normal mice⁷, but also prevented the formation of fibrous adhesions in the peritoneal cavities of these mice²⁵. Moreover, the application of drugs suppressing fibrin formation (heparin) or enhancing fibrinolysis (urokinase) again resulted in significantly enhanced CLP-induced mortality rates (B.E. *et al.*, manuscript in preparation). Together with histamine, mast-cell-derived TNF may contribute to fibrin formation by increasing vascular permeability, inducing tissue factor and inhibiting fibrinolysis²⁶.

In conclusion, our results show that mast cells and mast-cell-derived TNF contribute critically to the initiation of a cascade of host defence mechanisms aimed at localizing and confining a bacterial infection at the sites of acute inflammation. Clarification of the mechanisms involved in improving the survival of mice from experimental septic peritonitis may be important for the clinical management of postoperative peritonitis, which still causes a high mortality of 60–80% (ref. 27). □

Received 4 January; accepted 26 March 1996.

1. Galli, S. J. *New Engl. J. Med.* **328**, 257–265 (1993).
2. Schwartz, L. B. *Curr. Opin. Immun.* **6**, 91–97 (1994).
3. Galli, S. J. & Kitamura, Y. *Am. J. Pathol.* **127**, 191–198 (1987).
4. Wershil, B. K. & Galli, S. J. *Adv. exp. Med. Biol.* **347**, 39–54 (1994).
5. Nakano, T. *et al. J. exp. Med.* **162**, 1025–1043 (1985).
6. Nakano, T., Kanakura, Y., Asai, H. & Kitamura, Y. *J. Immunol.* **138**, 544–549 (1987).
7. Echtenacher, B., Falk, W., Männel, D. & Krammer, P. H. *J. Immunol.* **145**, 3762–3766 (1990).

8. Gordon, J. R., Burd, P. R. & Galli, S. J. *Immun. Today* **11**, 458–464 (1990).
9. Gordon, J. R. & Galli, S. J. *Nature* **346**, 274–276 (1990).
10. Gordon, J. R. & Galli, S. J. *J. exp. Med.* **174**, 103–107 (1991).
11. Zhang, Y., Ramos, B. F. & Jakschik, B. *Science* **258**, 1957–1959 (1992).
12. Leal-Berumen, I., Conlon, P. & Marshall, J. S. *J. Immunol.* **152**, 5468–5476 (1994).
13. Matsuda, H., Kawakita, K., Kiso, Y., Nakano, T. & Kitamura, Y. *J. Immunol.* **142**, 927–931 (1989).
14. Ansel, J. C., Brown, J. R., Payan, D. G. & Brown, M. A. *J. Immunol.* **150**, 4478–4485 (1993).
15. Yano, H., Wershil, B. K., Arizano, N. & Galli, S. J. *J. clin. Invest.* **84**, 1276–1286 (1989).
16. Sher, A., Hein, A., Moser, G. & Caulfield, J. P. *Lab. Invest.* **41**, 490–499 (1979).
17. Vranian, G. Jr, Conrad, D. H. & Ruddy, S. J. *Immunol.* **126**, 2302–2306 (1981).
18. Otani, I., Conrad, D. H., Carlo, J. R., Segal, D. M. & Ruddy, S. J. *Immunol.* **129**, 2109–2112 (1982).
19. Malaviya, R. *et al. J. Immunol.* **152**, 1907–1914 (1994).
20. Malaviya, R., Ross, E., Jakschik, B. & Abraham, S. N. *J. clin. Invest.* **93**, 1645–1653 (1994).
21. Wershil, B. K., Wang, Z.-S., Gordon, J. R. & Galli, S. J. *J. clin. Invest.* **87**, 446–453 (1991).
22. Tracey, K. J. *et al. Science* **234**, 470–474 (1986).
23. Rothstein, J. L. & Schreiber, H. *Proc. natn. Acad. Sci. U.S.A.* **85**, 607–611 (1988).
24. McClean, K. L., Sheehan, G. J. & Harding, G. K. *Clin. Infect. Dis.* **19**, 100–116 (1994).
25. Echtenacher, B., Falk, W., Männel, D. N. & Krammer, P. H. in *Host Defense Dysfunction in Trauma, Shock and Sepsis* (eds Faist, E., Meakins, J. L. & Schildberg, F. W.) 755–758 (Springer, Berlin, 1993).
26. Ramos, B. F., Zhang, Y. & Jakschik, B. A. *Eur. J. Immunol.* **22**, 2381–2385 (1992).
27. Holzheimer, R. G., Muhrer, K. H., L'Allemand, N., Schmidt, T. & Henneking, K. *Infection* **19**, 447–452 (1991).
28. Hültner, L. *et al. J. Immunol.* **142**, 3440–3446 (1989).
29. Egger, D. *et al. J. Immunol.* **154**, 1830–1837 (1995).
30. Alexander, H. R. *et al. J. clin. Invest.* **88**, 34–39 (1991).

ACKNOWLEDGEMENTS. B.E. and L.H. contributed equally to this work. We thank G. W. Bornkamm and S. J. Galli for discussion; G. Reisbach for FACS analyses; P. Dörmer and R. Mailhammer for critically reading the manuscript; H. Broszeit and B. Ruhland for technical assistance; and E. Ruppertsdorfer for maintaining the W/W^v breeding facility.

CORRESPONDENCE and requests for materials should be addressed to L.H. (e-mail: hultner@gsf.de).

Mast cell modulation of neutrophil influx and bacterial clearance at sites of infection through TNF- α

Ravi Malaviya*, Teruo Ikeda†, Elaine Ross* & Soman N. Abraham*‡

* Department of Pathology, Barnes-Jewish Hospital, and ‡ Departments of Pathology and Molecular Microbiology, Washington University School of Medicine, St Louis, Missouri 63110, USA

† Department of Veterinary Microbiology, Azabu University, Kanagawa, Japan

ALTHOUGH mast cells have been implicated in a variety of inflammatory conditions including immediate hypersensitivity and interstitial cystitis,^{1,2} their physiological role in the body is unknown. We investigated the role of mast cells in host defence against bacterial infections using a well characterized mast-cell-deficiency mouse model^{3,4}. We report here that mast cells, which are selectively located at portals of bacterial entry, are important to host defence. Mast-cell-deficient WBB6F1-W/W^v mice (W/W^v) were up to 20-fold less efficient in clearing enterobacteria than control WBB6F1+/+ (+/+) mice or mast-cell-reconstituted W/W^v (W/W^v + MC) mice. With higher bacterial inocula, only W/W^v mice died (80%). The limited bacterial clearance in W/W^v mice directly correlated with impaired neutrophil influx. The mast-cell chemoattractant TNF- α was implicated in the neutrophil response because TNF- α was locally released only in +/+ and W/W^v + MC mice, TNF- α -specific antibodies blocked over 70% of the neutrophil influx, and purified mast cells released TNF- α upon incubation with bacteria. Additionally, the type-1 fimbrial subunit, FimH, was the necessary enterobacterial component for mast-cell activation and neutrophil influx because an isogenic FimH⁻ mutant evoked a limited neutrophil response in +/+ mice compared to wild-type bacteria.

A mouse-virulent strain of *Klebsiella pneumoniae*, KPA1, was

injected into the peritoneal cavities of +/+ and W/W^v mice. After 48 h, the number of bacteria surviving in the peritoneal cavities of W/W^v mice was about 20-fold higher than in +/+ mice (Fig. 1a). Apparently, the mast-cell-deficiency impaired the animals' ability to clear the infectious agent. To confirm that the observed difference in bacterial clearance was solely due to mast cells but not other abnormalities, mast-cell-reconstituted W/W^v mice were challenged with the enterobacteria. The number of surviving bacteria in the peritoneal cavities of W/W^v + MC mice was comparable to that in +/+ mice (Fig. 1a). When 6-times-higher doses of KPA1 bacteria were injected into the three groups of mice, 80% of the W/W^v mice died within 24 h, whereas no deaths were observed in +/+ or W/W^v + MC mice. These data demonstrate the protective role of mast cells and reveal that the mechanism involves bacterial clearance.

Because mast cells display functional heterogeneity based on their location^{1,5}, their behaviour at a different and more natural site of infection was investigated. W/W^v, +/+ and W/W^v + MC mice were intranasally challenged with a pulmonary isolate of *K. pneumoniae*, KP1415, and the number of viable bacteria in the lungs of the mice was determined. In contrast to the two groups of mast-cell-sufficient mice, bacterial clearance in the lungs of W/W^v mice was significantly impaired (Fig. 1b). Microscopic examination of lung tissue sections from W/W^v and +/+ mice after bacterial challenge revealed remarkably high numbers of neutrophils in the latter group of mice, implying neutrophil involvement in bacterial clearance. To confirm this finding, we compared the number of neutrophils in the bronchoalveolar lavage (BAL) of W/W^v, +/+ and W/W^v + MC mice 6 h after intranasal challenge with *K. pneumoniae* KP1415. A significant decrease in myeloperoxidase, a specific neutrophil marker, was detected in the BAL of W/W^v compared to either the +/+ or W/W^v + MC mice (Fig. 1c). Thus, the reduced bacterial clearance in the lung of W/W^v relative to +/+ or W/W^v + MC mice directly correlates with reduced neutrophil influx in the mouse BAL.

Next, we sought to identify which bacterial component(s) was responsible for mast-cell activation leading to neutrophil influx and the mast cell factor(s) responsible for recruiting neutrophils. Our previous studies have implicated FimH, the minor subunit of type 1 fimbriae, as the determinant on enterobacteria responsible for promoting bacterial binding to mast cells^{6–8}. We compared the

neutrophil-recruiting abilities of a well characterized host strain, *Escherichia coli* ORN103, bearing either the recombinant plasmid pBP7, encoding wild-type *K. pneumoniae* type 1 fimbriae^{9,10}, or the plasmid pBP799, encoding FimH⁻ *K. pneumoniae* fimbriae¹¹. Because the FimH⁻ strain failed to invade the mouse lung, we used the infectious peritonitis model. The neutrophil influx in $+/+$ mice evoked by *E. coli* ORN103(pBP799) (5.0 ± 0.88 mU) was 6-fold lower than the influx induced by *E. coli* ORN103(pBP7) (32.9 ± 2.9 mU; $P < 0.01$). This finding implies that FimH, the fimbrial tip adhesin on *K. pneumoniae* type 1 fimbriae, is responsible for activating mast cells and is a potent recruiter of neutrophils *in vivo*. The radial arrangement of the fimbriae probably ensures that FimH is the first bacterial component sterically able to mediate productive interactions with mast cells and neutrophils.

Mast cells are capable of releasing several neutrophil chemoattractants, but the role of TNF- α is of particular interest because the mast cell is the only known cell to prestore TNF- α and is thus able to release this mediator immediately upon activation^{3,12,13}. A profile of extracellular TNF- α levels and the corresponding neutrophil population in the peritoneal fluids of control mice following instillation of type-1-fimbriated ORN103(pBP7) is shown in Fig. 2a and b, respectively. The initial influx of neutrophils into the peritoneum appears to coincide with a burst in extracellular TNF- α . Presumably because the activating agent is a bacterial lectin, the time course of TNF- α release is distinct from that described previously in an immune complex peritonitis model³. It is not known why the TNF- α levels in the peritoneum dropped after the 1 h time point in spite of the influx of a sizeable number of neutrophils.

To show that the source of the TNF- α burst was the mast cell, we compared TNF- α levels in the peritoneal cavities of W/W^v,

$+/+$ and W/W^v + MC mice after challenge with type-1-fimbriated ORN103(pBP7). The TNF- α levels in W/W^v mice were substantially lower than in either $+/+$ or W/W^v + MC mice (Fig. 3), confirming that the cytokine originated from the mast cells. The residual levels of TNF- α seen in W/W^v mice could be ascribed to macrophages and other peritoneal cells¹⁴. Additionally, when monolayers of cultured mast cells were exposed to 1×10^7 type-1-fimbriated bacteria for 1 h, significant amounts of TNF- α (2.8 ± 1.2 ng per 10^6 mast cells) were released, indicating the ability of mast cells to release TNF- α in response to bacteria. Thus, mast cells may be an important source of the elevated local and systemic TNF- α levels typically seen following Gram-negative bacterial infections^{15,16}.

To demonstrate the role of mast-cell-derived TNF- α in neutrophil recruitment, the inhibitory effect of antibodies to TNF- α on neutrophil influx in W/W^v, $+/+$ and W/W^v + MC mice was investigated. Specific monoclonal antibodies to TNF- α , but not to interleukin-1 β (control), blocked up to 70% of the neutrophil influx into the peritoneal cavities of $+/+$ and W/W^v + MC mice (Fig. 4). Notably, TNF- α is not the sole mast-cell modulator of neutrophil influx. We have determined that two other mast-cell chemoattractants, leukotriene B₄ and interleukin-8, play a significant role in neutrophil influx following bacterial challenge (data not shown). The relevance of these mast-cell mediators is also indicated by the fact that neutralization of the effect of TNF- α in mast-cell-sufficient mice, by administration of anti-TNF- α antibodies¹⁷, did not increase their mortality (in comparison to mice treated to control antibody) when they were peritoneally challenged with 6×10^7 of the virulent KPA1 bacteria.

Mast cells have the capacity to recognize bacterial surface molecules specifically and, through the release of neutrophil

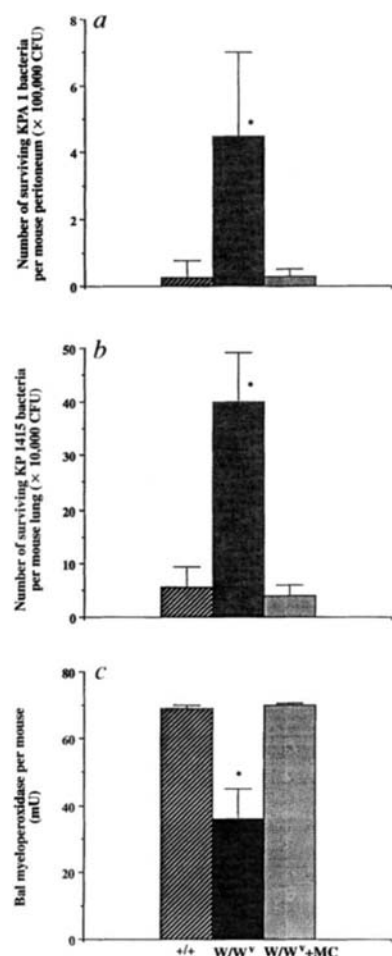


FIG. 1 Effect of mast-cell deficiency on bacterial clearance and the correlation of neutrophil influx with bacterial clearance in mast-cell sufficient mice. a, Clearance of *K. pneumoniae* KPA1 in the peritoneal cavities and b, *K. pneumoniae* KP1415 in the lungs of mast-cell deficient and sufficient mice. c, Corresponding neutrophil population in the BAL of mice challenged with strain KP1415. The data are expressed as mean $\pm$ s.e.m. ($n =$ at least 6 mice); * $P < 0.05$, on comparison of W/W^v with $+/+$ or W/W^v + MC mice as determined by Wilcoxon Signed Rank test. The protocol used in the care and treatment of mice was in accordance with institutional guidelines.

METHODS. Ninety-day-old normal WBB6F1 $+/+$ ($+/+$) and mast-cell-deficient WBB6F1-W/W^v (W/W^v) mice were purchased from Jackson Laboratories (Bar Harbor, ME). To generate W/W^v + MC mice for the peritonitis model, adoptive transfer of mast cells (> 98% pure) cultured from the bone marrow of $+/+$ mice⁶, into W/W^v mice was achieved by intraperitoneal injection of 1.0×10^7 mast cells³. The adoptive transfer of exogenous mast cells corrects only the mast-cell deficiency in W/W^v mice and no other deficiency^{18,19}. At time of bacterial challenge in $+/+$ mice there were 0.6×10^6 mast cells per peritoneal cavity of which 10% stained with alcian blue, but not safranin, and 90% stained with safranin. The peritoneal cavities of W/W^v + MC mice contained 0.8×10^6 mast cells of which 10% stained alcian blue, 5% stained with safranin and 85% exhibited staining with both dyes. For the pneumonitis model, adoptive transfer of mast cells into the lungs of W/W^v mice was achieved by tail-vein injection of 5.0×10^6 mast cells²⁰. The range of mast cells per ultrathin sagittal section of lung was 2–3 per section in W/W^v + MC mice as opposed to 1–3 per section in $+/+$ mice. In the $+/+$ mice, the mast cells stained either with alcian blue or safranin whereas in W/W^v + MC mice, the mast cells stained either with alcian blue or with both alcian blue and safranin. Although, according to staining with alcian blue and safranin, the mast cell population in W/W^v + MC mice was not identical to that in $+/+$ mice, previous studies demonstrated that full restoration of some mast-cell functions was achieved in these models at 3 weeks after mast-cell reconstitution³. Haematocrit values in W/W^v + MC mice ranged from 33–36%. This compares with 36–43% in $+/+$ mice and 34–36% in W/W^v mice. Infectious peritonitis: All three groups of mice were injected i.p. with 1.0×10^7 of a virulent strain of *K. pneumoniae*, KPA1²¹. After 48 h of incubation, all the mice were killed and their peritoneal cavities lavaged with 2 ml sterile PBS (pH 7.4). The peritoneal lavage from each group of mice was assayed for bacterial viability by a standard plate method⁶. No viable bacteria was detected in the peritoneal lavage from mice after injection of vehicle (PBS) alone. Infectious pneumonitis: Intranasal instillation of anaesthetized W/W^v, $+/+$ and W/W^v + MC mice with KP1415 was achieved with the help of a 26-gauge blunt needle (1×10^8 bacteria in 50 μ l suspension). After 6 h, each group of mice was killed and 5 ml bronchioalveolar lavage (BAL) was collected with 10 aliquots (0.5 ml) of sterile saline²². Neutrophil population in the lavage was evaluated by measurement of myeloperoxidase activity by colorimetry of cell lysates²³. In addition, the left lung from each mouse was aseptically removed, weighed and homogenized. The number of viable bacteria in the left lung was determined.

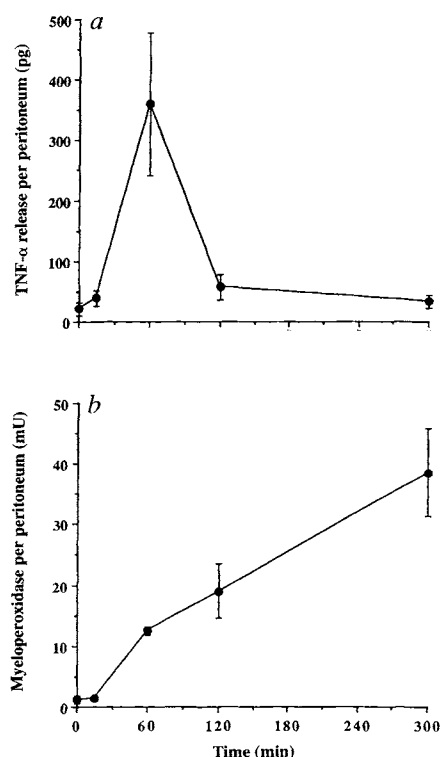


FIG. 2 Time course of TNF- α release (a) and neutrophil influx (b) in the peritoneal cavities of mice following bacterial instillation. The data are expressed as mean $\pm$ s.e.m. The zero time point represents measurements from PBS-treated mice.

METHODS. A large group of mast-cell sufficient mice were challenged i.p. with 2.5×10^7 ORN103(pBP7) expressing *K. pneumoniae* type 1 fimbriae and at 15 min, 1, 2 and 5 h intervals thereafter, mice were killed and their peritoneal cavities lavaged and assayed for extracellular TNF- α and for neutrophils. TNF- α was assayed by a standard cytotoxicity assay²⁴ and neutrophils were assayed as described in Fig. 1. Monoclonal antibodies to TNF- α completely blocked the observed cytotoxicity in the TNF- α assay confirming that TNF- α , and not any bacterial product, was responsible for cytotoxicity to the indicator cells²⁴.

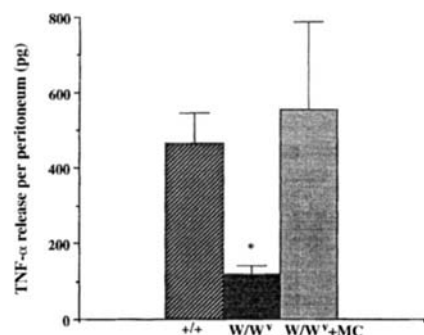


FIG. 3 Extracellular TNF- α in the peritoneal cavities of mast-cell deficient and sufficient mice following bacterial instillation. The data are expressed as mean $\pm$ s.e.m. ($n = 6$ mice). * $P < 0.05$ on comparison of W/W⁻ with +/+ or W/W⁻ + MC mice.

METHODS. As described previously, 2.5×10^7 of type-1-fimbriated ORN103(pBP7) bacteria were injected i.p. into W/W⁻, +/+ and W/W⁻ + MC mice. After 1 h (the time point that corresponded to the TNF- α burst in Fig. 2), the peritoneal cavities of the mice were lavaged and assayed for extracellular TNF- α .

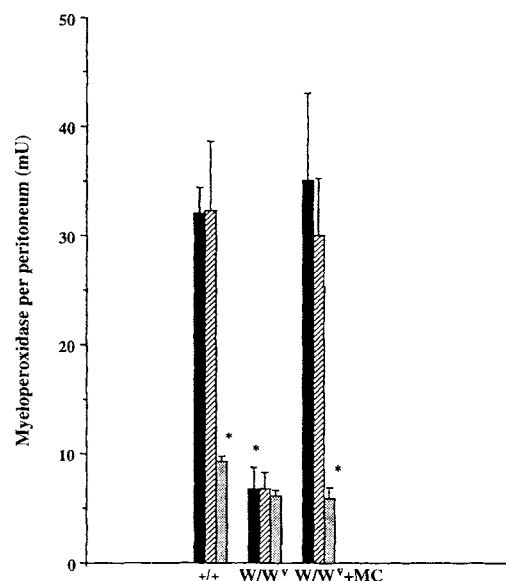


FIG. 4 Effect of TNF- α -specific antibody on the neutrophil response following intraperitoneal challenge with type 1 fimbriated enterobacteria. The data are expressed as mean $\pm$ s.e.m. ($n = 6-9$ mice); * $P < 0.05$, on comparison of neutrophil influx between W/W⁻ and +/+ or W/W⁻ + MC mice, and on comparison of neutrophil influx between TNF- α antibody-treated (shaded bars) and control antibody-treated (hatched bars) +/+ or W/W⁻ + MC mice. Black bars, no pretreatment.

METHODS. Mast-cell-deficient (W/W⁻), normal (+/+), and mast-cell reconstituted (W/W⁻ + MC) mice were injected i.p. with 2.5×10^7 *E. coli* ORN103(pBP7) expressing *K. pneumoniae* type 1 fimbriae^{9,10}. After 5 h of incubation, all the mice were killed and their peritoneal cavities lavaged as described in Fig. 1. To examine the specific involvement of TNF- α in neutrophil elicitation, W/W⁻, +/+ and W/W⁻ + MC mice were injected i.p. with a neutralizing monoclonal antibody to mouse TNF- α (250 μ g per mouse) or control monoclonal antibody (specific for mouse IL-1 β) (250 μ g per mouse) 18 h before challenge³. Mice that did not receive any antibody (untreated) were given PBS intraperitoneally. Myeloperoxidase activity in peritoneal cavities of PBS-treated mice was 0.17 mU per peritoneum.

chemoattractants, modulate neutrophil influx and bacterial clearance. This represents a pivotal, logical, but hitherto unconfirmed, physiological role for a host cell that is replete with pharmacologically active compounds and which is strategically stationed at the host-environment interface. □

Received 21 November 1995; accepted 4 March 1996.

- Galili, S. J. *New Engl. J. Med.* **328**, 257-263 (1993).
- Malone, D. G. & Metcalfe, D. D. *Ann. Allergy* **61**, 27-30 (1988).
- Zhang, Y., Ramos, B. F. & Jakschik, B. A. *Science* **258**, 1957-1959 (1992).
- Kitamura, Y., Go, S. & Hatanaka, K. *Blood* **52**, 447-452 (1978).
- Bienenstock, J., Befus, A. D., Pearce, F., Denburg, J. & Goodacre, R. J. *Allergy clin. Immun.* **70**, 407-412 (1982).
- Malaviya, R. et al. *J. Immun.* **152**, 1907-1914 (1994).
- Malaviya, R., Ross, E. A., Jakschik, B. A. & Abraham, S. N. *J. clin. Invest.* **93**, 1645-1653 (1994).
- Malaviya, R., Twisten, N. J., Ross, E. A., Abraham, S. N. & Pfeifer, J. D. *J. Immun.* **156**, 1490-1496 (1996).
- Gerlach, C. F., Clegg, S. & Allen, B. J. *Bact.* **171**, 1262-1270 (1989).
- Abraham, S. N., Sun, D., Dale, J. B. & Beachy, E. H. *Nature* **336**, 682-684 (1988).
- Madison, B., Ofek, I., Clegg, S. & Abraham, S. N. *Infect. Immun.* **62**, 843-848 (1994).
- Gordon, J. R. & Galli, S. J. *J. exp. Med.* **174**, 103-117 (1991).
- Wershil, B. K., Wang, Z.-H., Gordon, J. R. & Galli, S. J. *J. clin. Invest.* **87**, 446-453 (1991).
- Vassalli, P. A. *Rev. Immun.* **10**, 411-452 (1992).
- Natanson, C., Hoffman, W. D., Suffredini, A. F., Eichacker, P. Q. & Danner, R. L. A. *Intern. Med.* **120**, 771-783 (1994).
- Arnold, J. W. et al. *Microbiol. Pathogen.* **14**, 217-227 (1993).
- Echtenacher, B., Falk, W., Mannel, D. N. & Krammer, P. H. *J. Immun.* **145**, 3762-3766 (1990).
- Kitamura, Y. et al. *J. exp. Med.* **150**, 482-490 (1979).
- Nakano, T. et al. *J. exp. Med.* **162**, 1025-1043 (1985).
- Suzuki, N. et al. *Am. J. Resp. cell Molec. Biol.* **9**, 475-483 (1993).
- Ofek, I. et al. *Infect. Immun.* **61**, 4208-4216 (1993).

22. Esposito, A. L., Quinn, L. A., Lucey, E. C. & Snider, G. L. *Am. Rev. respir. Dis.* **135**, 676–688 (1987).
 23. Bradley, P. P., Priebat, D. A., Christensen, R. D. & Rothstein, G. J. *Invest. Dermatol.* **78**, 206–209 (1982).
 24. Sheehan, K. C. F., Ruddle, N. H. & Schreiber, R. D. J. *Immun.* **142**, 3884–3893 (1989).

ACKNOWLEDGEMENTS. We thank B. Jakschik for her helpful suggestions and comments, F. Ross, J. Pfeifer, R. Lorenz and D. Baorto, all of Washington University, St Louis, are gratefully acknowledged for critical review of this manuscript. We thank I. Ofek (University of Tel Aviv, Israel) for donation of *K. pneumoniae* KPA1, S. Clegg (University of Iowa, Iowa City) for donation of plasmids pBP7 and pBP799, and K. C. F. Sheehan and R. Schreiber (Department of Pathology, Washington University, St Louis) for providing monoclonal antibodies to TNF- α . Finally, we would like to acknowledge Marshall Plaut at the National Institutes of Health for his support. Research funds came from the National Institutes of Health, the Interstitial Cystitis Association of America and from a Monsanto-Searle-Washington University Biomedical Research Agreement.

CORRESPONDENCE and requests for materials should be addressed to S.N.A. (e-mail abralab@pathology.wustl.edu).

Specific phosphorylation of SR proteins by mammalian DNA topoisomerase I

Ferdinand Rossi*, Emmanuel Labourier*, Thierry Forné*, Gilles Divita†, Jean Derancourt‡, Jean François Riou‡, Etienne Antoine*, Guy Cathala*, Claude Brunel* & Jamal Tazi*

* Institut de Génétique Moléculaire (IGM), UMR 5535 CNRS, Université de Montpellier II, † Centre de Recherche de Biochimie Macromoléculaire, CNRS LP 9008–INSERM Unité 249, CNRS BP 5051, route de Mende, F34033 Montpellier CEDEX 1, France

‡ Rhône-Poulenc Rorer SA, Centre de Recherche de Vitry-Alfortville, Unité de Cancérologie, CRVA 14, Département Biologie, Quai Jules Guesde, 94403 Vitry-sur-Seine, France

SEVERAL metazoan splicing factors are characterized by ribonucleoprotein (RNP) consensus sequences and arginine-serine repeats (RS domain)^{1–8} which are essential for their function in splicing^{9,10}. These include members of the SR-protein family (SC35, SF2/ASF), the U1 small nuclear (sn)RNP protein (U1-70K) and the U2 snRNP auxiliary factor (U2AF). SR proteins are phosphorylated *in vivo*¹¹ and the phosphorylation state of U1-70K's RS domain influences its splicing activity¹². Here we report the purification of a protein kinase that is specific for SR proteins and show that it is DNA topoisomerase I. This enzyme lacks a canonical ATP-binding motif but binds ATP with a dissociation constant of 50 nM. Camptothecin and derivatives, known to be specific inhibitors of DNA topoisomerase I, strongly inhibit the kinase activity in the presence of DNA and affect the phosphorylation state of SR proteins. Thus, DNA topoisomerase I may well be one of the SR protein kinases operating *in vivo*.

The identification of a U1 snRNP-associated protein kinase that selectively phosphorylates the RS domain of SR proteins¹³ led us to search for protein(s) that have this activity in HeLa cell nuclear extract, which has been used as a source to prepare U snRNPs (ref. 14). In addition to SF2/ASF, the predominant phosphorylated protein, other endogenous proteins were also phosphorylated by the HeLa extract (Fig. 1a, lane 1). Specific SF2/ASF phosphorylation was purified by two rounds of ammonium sulphate precipitation, followed by fractionation on a nickel-resin column and elution with imidazole (Fig. 1a, lane 6). Phosphoaminoacid and phosphopeptide analyses demonstrated that phosphorylation of SF2/ASF by this fraction was exclusively at serine residues that map to the RS domain (data not shown). As recombinant SC35 protein and gel-purified individual SR proteins derived from calf thymus⁶ were also phosphorylated with this kinase preparation but histone H1 or casein were not (data not shown), we concluded that we had isolated a specific SR-protein kinase.

This kinase activity comigrated with an activity with the ability to relax negatively supercoiled DNA (Fig. 1b) in an assay for DNA topoisomerase I. Analysis of proteins by SDS-PAGE and Coomassie blue staining showed that the single band associated with these activities migrated with an apparent relative molecular mass (M_r) of 100,000 (100K; Fig. 1c). In a two-dimensional gel¹⁵, this fraction ran as a smear of M_r 100K and with a pI ranging from 8.0 to 8.3 (Fig. 1d), which generated a single spot focusing at pH 8.3 following alkaline phosphatase treatment (data not shown). Thus the fraction contains a single 100K phosphoprotein with a range of pIs due to different levels of phosphorylation. Peptide sequencing established that this 100K protein is human DNA topoisomerase I. The pronounced affinity for nickel-resin of DNA topoisomerase I is likely to be conferred by several histidine residues clustered at its amino terminus¹⁶.

Human DNA topoisomerase I was expressed at high levels in a baculovirus system¹⁷ and purified to apparent homogeneity as already described. It had the same apparent M_r as the protein from HeLa cells (Fig. 2a), was effective in phosphorylating recombinant SF2/ASF (Fig. 2b), having a comparable specific activity to HeLa cell DNA topoisomerase I (2.5 and 3 nmol phosphate transferred per min per mg of protein, respectively), and efficiently relaxed supercoiled DNA in an ATP-independent manner (Fig. 2c). This preparation may still be contaminated with trace amounts of 135K insect DNA topoisomerase I because similarly purified preparations from Sf21 cells infected with virus encoding β -galactosidase (BacPAK6) had slight protein kinase and DNA topoisomerase I activities (Fig. 2b, c). However, serial dilution of both preparations indicated that at a 1/10 dilution protein kinase and DNA topoisomerase I activities from control cells became undetectable, whereas purified recombinant DNA topoisomerase I was still active (Fig. 2b, c); this indicates that the overexpressed and purified recombinant human DNA topoisomerase I also acts as a SR protein kinase.

DNA topoisomerase I lacks any canonical ATP-binding site like that found in all other protein kinases identified so far. We therefore investigated the ATP-binding properties of DNA topoisomerase I. Human DNA topoisomerase I contains 13 tryptophan residues which can be used as a probe for monitoring the kinetics of enzyme-substrate interactions. Addition of ATP to both HeLa or recombinant DNA topoisomerase I induced a marked quenching of the intrinsic fluorescence (22 and 26%, respectively; Fig. 3), indicating that some tryptophan residues are directly or indirectly involved in nucleotide binding. Curve fitting was consistent with DNA topoisomerase I having one ATP-binding site with a dissociation constant for ATP binding of ~ 50 nM (48 nM and 55 nM for HeLa and recombinant DNA topoisomerase I, respectively). Similar values (50 nM) were obtained when binding was monitored by quenching of 2', (3')-N-methyl-anthraniloyl-ATP fluorescence (data not shown). DNA topoisomerase I is thus the major ATP-binding component in our preparations and contains a single high-affinity binding site.

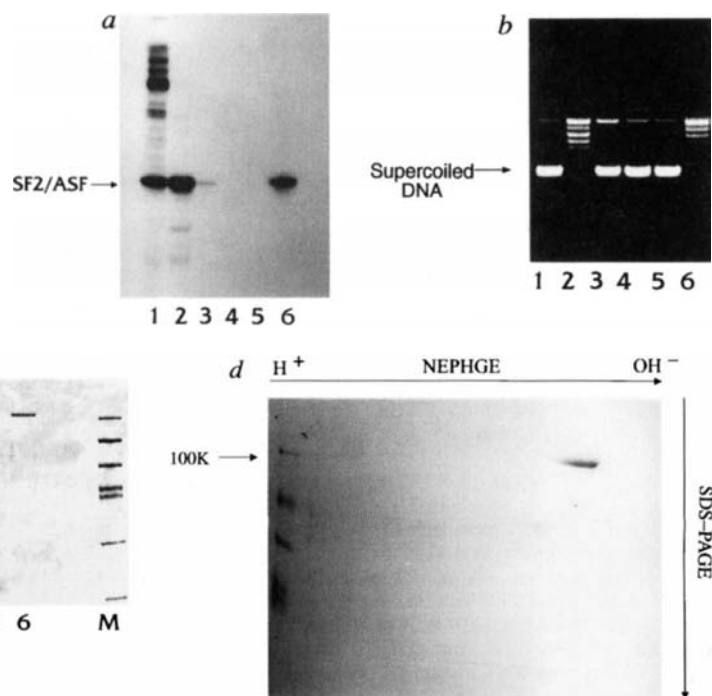
DNA topoisomerase I is a particular target of the antitumour alkaloid camptothecin^{18–20}. Purified SR proteins metabolically labelled with ³²P from HeLa cells treated with topotecan (a water-soluble derivative of camptothecin) had the same protein pattern as that from untreated cells (Fig. 4a), showing that the drug affected neither the purification profile nor the overall phosphorylation level of these SR proteins. Two-dimensional gel analysis of these SR proteins revealed that each protein is made of several isoelectric variants (Fig. 4b, c) which vary chiefly as a result of their phosphorylation state, being sensitive to alkaline phosphatase treatment (data not shown). Topotecan treatment abolished the labelling of the two most phosphorylated variants of SRp75 and strongly diminished those of SRp30, SRp40 and SRp55 while enhancing the radioactivity contained in the less phosphorylated variants (Fig. 4c). When topotecan was removed from the culture medium and cells grown for a further three hours in the presence of cold inorganic phosphate, the labelling pattern returned to that of untreated cells (data not

shown). Treatment of cells with camptothecin instead of topotecan gave similar results (not shown). As camptothecin derivatives did not completely inhibit SR protein phosphorylation, other protein kinase(s) not affected by camptothecins are also likely to be active *in vivo*, although at different sites from DNA topoisomerase I^{21,22}.

In vitro kinase assays in the presence of increasing concentrations of camptothecin (up to 500 μ M) showed that this drug itself

has no effect on kinase activity in the absence of DNA (only one concentration is shown in Fig. 4d). Similarly, DNA alone had only a slight inhibitory effect on protein kinase activity, but camptothecin and DNA together had a much stronger effect (Fig. 4d). Camptothecin stabilizes a ternary complex in which topoisomerase I is covalently bound to DNA, so a combination of camptothecin and DNA is needed to prevent effectively the protein kinase activity of DNA topoisomerase I.

FIG. 1 Purification of SR protein kinase *a*, Autoradiogram of kinase assays carried out with nuclear extract (lane 1), ammonium sulphate precipitate (lane 2), unbound protein to nickel resin (lane 3), first and second washes of the resin (lane 4 and 5, respectively) and the 40 mM imidazole eluate (lane 6). *b*, DNA topoisomerase I activities of the same fractions as in *a*, apart from lane 1, which was assayed in buffer B (see Methods). *c*, SDS-PAGE analysis of 10 μ l nuclear extract (lane 1), 15 μ l ammonium sulphate precipitate (lane 2), 30 μ l proteins unbound to nickel resin (lane 3), 30 μ l first wash (lane 4), 30 μ l second wash (lane 5) and 30 μ l of 40 mM imidazole eluate (lane 6) on a 10% gel stained with Coomassie blue. Lanes M, size markers. *d*, Two-dimensional gel electrophoresis of 40 mM imidazole eluate fraction; Nonequilibrium pH gradient electrophoresis. **METHODS.** Nuclear extracts were prepared from isolated nuclei as described^{14,27}, except that they were dialysed against buffer A (10 mM triethanolamine, pH 7.9, 100 mM KCl, 0.2 mM EDTA, 0.5 mM DTT) before storage at -80°C . Nuclear extract (500 μ l, 750 μ g protein) was clarified by centrifugation (4 min, 17,000g), mixed with an equal volume of 2 M ammonium sulphate (4 $^{\circ}\text{C}$, 1 h), clarified (10 min, 17,000g), precipitated with 248 mg ammonium sulphate (4 $^{\circ}\text{C}$, 1 h) and collected by centrifugation. The pellet (250 μ g protein) was resuspended (in 500 μ l of buffer B: 50 mM HEPES, pH 7.0, 10 mM MgCl_2 , 3 mM MnCl_2 , 50 mM KCl, 0.5 mM DTT), incubated with 50 μ l nickel-agarose beads (QIAGEN, 30 min, 4 $^{\circ}\text{C}$). After two washes (in 300 μ l buffer B), bound kinase activity was eluted with 40 mM imidazole. Protein concentrations were typically 10–20 $\mu\text{g ml}^{-1}$ (from Coomassie blue staining, using BSA as a standard). For protein kinase assays, the test fraction (10 μ l), recombinant ASF/SF2 protein (300 ng in buffer B) and [γ -³²P]ATP (3 μCi , 3,000 Ci mmol⁻¹) were incubated (12 μ l final volume, 30 $^{\circ}\text{C}$, 30 min), mixed with 3 $\times$ Laemmli buffer (5 μ l) and loaded onto 10% SDS-polyacrylamide gels. Fractionated phosphoproteins were revealed by autoradiography. Gel slices containing radioactive SF2/ASF substrate were excised and counted for ³²P in a liquid scintillation counter giving a specific activity for DNA topoisomerase I kinase activity of 3 nmol phosphate transferred per min per μg under standard kinase reaction conditions. Kinase reactions with 100 μM ATP and 1 μg DNA topoisomerase I allow phosphorylation of 10 pmol SF2/ASF, showing that DNA topoisomerase I transfers three phosphates per molecule of SF2/ASF. To assay for DNA



topoisomerase I activity, test fractions (2 μ l) were incubated with supercoiled DNA plasmid (500 ng in buffer B, 10 μ l final volume, for 30 min at 30 $^{\circ}\text{C}$). DNA plasmids were extracted (phenol-chloroform), precipitated (ethanol) and electrophoresed (0.8% agarose gel in TBE buffer). DNA topoisomers were revealed by ethidium bromide staining. For 2D gel analysis, 40 mM imidazole eluate fractions (1 ml) were incubated with (not shown) or without calf intestinal alkaline phosphatase (GIBCO BRL; 12 units, 30 $^{\circ}\text{C}$, 1 h), precipitated (10% trichloroacetic acid), washed (2 $\times$ cold acetone), resuspended in loading buffer and subjected to 2D gel electrophoresis¹⁵: first dimension, pH; second dimension, 10% SDS-PAGE. Gels were stained with Coomassie blue after soaking in trichloroacetic acid (5%, for 30 min). SF2/ASF was expressed in TG1 bacteria transfected with plasmid containing ASF-1 cDNA (gift from J. Manley) and purified from inclusion bodies¹. SC35 protein expressed in a baculovirus system and purified⁴ was kindly provided by J. Stevenin and R. Gattoni. SR proteins were purified from calf thymus⁶.

FIG. 2 Overexpression and purification of recombinant topoisomerase I. *a*, Analysis of purified protein fractions from HeLa cells (lane 1), Sf21 cells infected with AcMNPV/hTOPI (lane 2) or BacPAK6 (control, lane 3). *b*, SR protein kinase assays (as for Fig. 1a) with 1 μ l crude (lane 1) or 1/10 diluted (lane 2) control fraction; 1 μ l crude (lane 3) or 1/10 (lane 4) baculovirus recombinant human DNA topoisomerase I fraction. *c*, DNA topoisomerase I assays (as for Fig. 1b), with lanes 1, 2 and 4, 5 corresponding to the fractions in lanes 1, 2 and 3, 4 in *b*; lane 3, 10 μ l buffer B. **METHODS.** Monolayer cultures of Sf21 cells were infected with AcMNPV/hTOPI (a generous gift from A. M. Zhelkovsky, Tufts University, Boston), or BacPAK6 (CLONTECH) at a multiplicity of infection of 10 and harvested by centrifugation three days post-infection. Nuclear extracts were prepared^{14,27} and recombinant DNA topoisomerase I purified as in Fig. 1, yielding approximately 4 mg DNA topoisomerase I from 10⁹ Sf21 cells infected with AcMNPV/hTOPI. Kinase specific activity of the purified recombinant human topoisomerase I was determined as in Fig. 1.

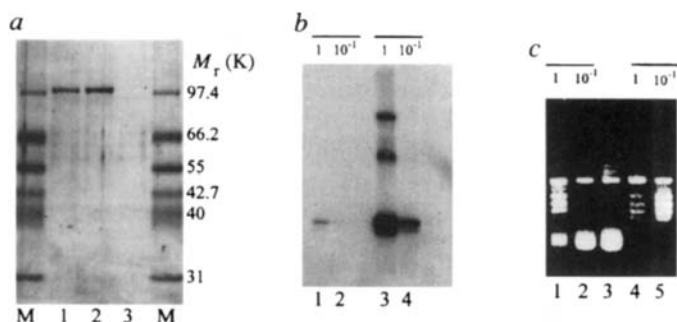


FIG. 3 Titration of ATP binding to DNA topoisomerase I using tryptophan fluorescence (arbitrary units) as a signal of binding. Fluorescence experiments were performed using a Fluorolog-II (Jobin Yvon) spectrofluorimeter at 25 °C. The enzymes (25 nM) were incubated in a standard buffer (50 mM Tris-HCl, 2 mM MgCl₂, 50 mM KCl). Binding of ATP was monitored by the quenching of the intrinsic tryptophan fluorescence of HeLa-cell purified (filled circles) and baculovirus recombinant (open circles) DNA topoisomerase I at 332 nm, on excitation at 295 nm. The decrease in protein fluorescence was fitted to the appropriate form of the quadratic equation^{28,29}. Best fits were obtained with K_d values of 48 and 55 nM and a maximal fluorescence quenching of 23 and 26% for HeLa-purified and recombinant human DNA topoisomerase I, respectively.

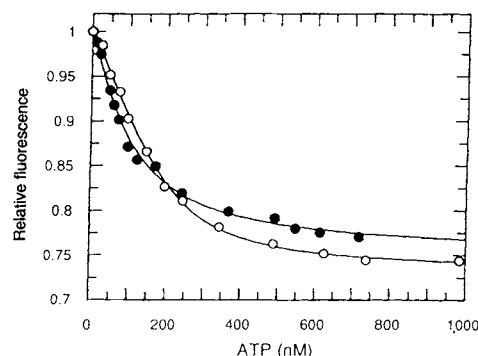
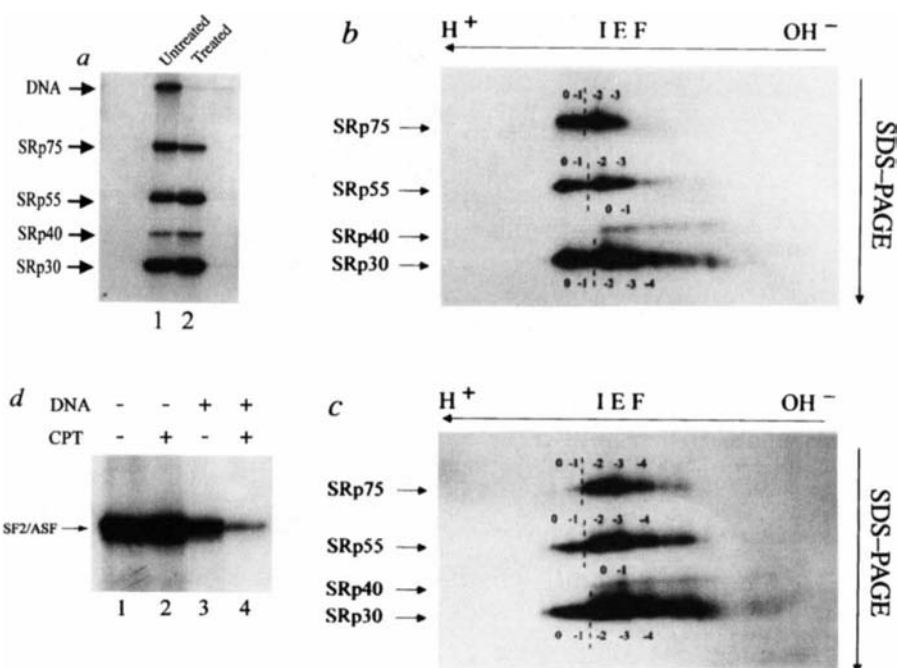


FIG. 4 Effect of topotecan on SR protein phosphorylation. **a**, A quarter of the total metabolically ³²P-labelled SR proteins, isolated from untreated (lane 1) or topotecan-treated (lane 2) HeLa cells, was analysed by 10% SDS-PAGE and detected by autoradiography. Equal amounts of protein were loaded in each lane (by Coomassie blue staining and immunoblot with mAb 104). In **b** and **c**, similar samples to those in **a** were analysed by two-dimensional gel electrophoresis³⁰. ³²P-labelled SR protein variants were isolated from **b**, untreated, or **c**, topotecan-treated cells and detected by autoradiography. The electrophoretic mobility of each variant was determined by comparison to mobilities of cold hnRNP proteins analysed in the same gel and stained with Coomassie blue. Assuming that one phosphate is responsible for the difference in mobility of two contiguous variants from each SR protein, we arbitrarily attributed the value zero to the most acidic variant and increasing negative numbers to preceding variants as they lost phosphates. Vertical bars indicate the position of hyperphosphorylated variants (0 and -1) relative to other less phosphorylated variants. **d**, Kinase assays were done as for Fig. 1, using 10 µl 40 mM imidazole eluate as a source of DNA topoisomerase I (lane 1), except that 500 µM camptothecin (lane 2), 100 nM plasmid DNA (lane 3) or 500 µM camptothecin and 100 nM plasmid DNA (lane 4) were added. METHODS. HeLa cells were resuspended (5×10^5 cells ml⁻¹) in phosphate-free minimum essential medium (2 litres) supplemented with fetal



calf serum (5%) and exposed to carrier-free H₃³²PO₄ (Amersham, 5 mCi, 12 h). Treated cells were grown under the same conditions except that 10 µM topotecan was included. SR proteins were immediately prepared from fresh labelled cells⁶

We have identified DNA topoisomerase I as a potential SR protein kinase in the cell. In contrast to other, well-characterized SR protein kinases^{21,22}, this protein has no obvious kinase domain but binds ATP with a K_d of 50 nM, a value consistent with that of known protein kinases, and phosphorylates specific serine residues in the RS domain. Despite considerable evidence for the involvement of DNA topoisomerase I in many processes of DNA metabolism²³⁻²⁶, its precise role is still uncertain. The discovery that DNA topoisomerase I has a protein phosphotransferase activity should lead to a better understanding of its biological function. □

Received 14 September 1995; accepted 1 March 1996.

- Ge, H., Zuo, P. & Manley, J. L. *Cell* **66**, 373-382 (1991).
- Krainer, A. R., Mayeda, A., Kozak, D. & Binns, G. *Cell* **66**, 383-394 (1991).
- Vellard, M., Sureau, A., Soret, J., Martinier, C. & Perbal, B. *Proc. natn. Acad. Sci. U.S.A.* **89**, 2511-2515 (1992).
- Fu, X. D. & Maniatis, T. *Science* **256**, 535-538 (1992).
- Zamore, P. D., Patton, J. G. & Green, M. R. *Nature* **355**, 609-614 (1992).
- Zahler, A. M., Lane, W. S., Stolk, J. A. & Roth, M. B. *Genes Dev.* **6**, 837-847 (1992).
- Moore, M. J., Query, C. C. & Sharp, P. A. in *RNA World* (eds Gestland, R. & Atkins, J.) 303-357 (Cold Spring Harbor Laboratory Press, New York, 1993).
- Fu, X. D. *RNA* **1**, 663-680 (1995).
- Zuo, P. & Manley, J. L. *EMBO J.* **12**, 4727-4737 (1993).
- Càceres, J. F. & Krainer, A. R. *EMBO J.* **12**, 4715-4726 (1993).

- Roth, M. B., Zahler, A. M. & Stolk, J. A. *J. Cell Biol.* **115**, 587-596 (1991).
- Tazi, J. et al. *Nature* **363**, 283-286 (1993).
- Wopmann, A. et al. *Nucleic Acids Res.* **21**, 2815-2822 (1993).
- Bach, M., Bringmann, P. & Lüthmann, R. *Meth. Enzym.* **181**, 232-256 (1990).
- O'Farrell, P. Z., Goodman, H. M. & O'Farrell, P. H. *Cell* **12**, 133-142 (1977).
- D'Arpa, P. et al. *Proc. natn. Acad. Sci. U.S.A.* **85**, 2543-2547 (1988).
- Zhelkovsky, A. M. & Moore, C. L. *Protein Expression and Purification* **5**, 364-370 (1994).
- Bjornsti, M. A., Benedetti, P., Vigilanti, G. A. & Wang, J. C. *Cancer Res.* **49**, 6318-6323 (1989).
- Hsiang, Y. H., Herizberg, R., Hecht, S. & Liu, L. T. *J. Biol. Chem.* **266**, 14873-14878 (1991).
- Froelich-Armon, S. J. & Osheroff, N. *J. Biol. Chem.* **270**, 21429-21432 (1995).
- Gui, J. F., Lane, W. S. & Fu, X. D. *Nature* **369**, 678-682 (1994).
- Colwill, K. et al. *EMBO J.* **15**, 265-275 (1996).
- Wang, J. C. A. *Rev. Biochem.* **54**, 665-697 (1985).
- Lee, P. L., Brown, S. D., Chen, A. & Hsieh, T. H. *Proc. natn. Acad. Sci. U.S.A.* **90**, 6656-6660 (1993).
- Merino, A., Madden, K. R., Lane, W. S., Champoux, J. J. & Reinberg, D. *Nature* **365**, 227-232 (1993).
- Gupta, M., Fujimori, A. & Pommier, Y. *Biochim. biophys. Acta* **1262**, 1-14, (1995).
- Zieve, G. W. & Penman, S. *J. molec. Biol.* **145**, 501-523 (1981).
- Divita, G., Goody, R. S., Gautheron, D. C. & Di Pietro, A. *J. Biol. Chem.* **268**, 13178-13186 (1993).
- Rittinger, K., Divita, G. & Goody, R. S. *Proc. natn. Acad. Sci. U.S.A.* **92**, 8046-8049 (1995).
- O'Farrell, P. H. *J. Biol. Chem.* **250**, 4007-4021 (1975).

ACKNOWLEDGEMENTS. We thank M. Silhol for technical assistance and P. Jeanteur for discussion and for comments on the manuscript. This work was supported by grants from the Association pour la Recherche Contre le Cancer to C.B. and J.T. F.R. was supported by a fellowship from the ligue nationale contre le cancer.

CORRESPONDENCE and requests for materials should be addressed to J.T. (e-mail: tazi@jones.igm.cnrs-mop.fr).

Adaptive inducibility of CREM as transcriptional memory of circadian rhythms

Nicholas S. Foulkes, Gilles Duval
& Paolo Sassone-Corsi

Institut de Génétique et de Biologie Moléculaire et Cellulaire,
1 rue Laurent Fries, 67404 Illkirch Cédex, CU de Strasbourg, France

THE CREM gene¹⁻⁴ encodes the transcriptional repressor ICER^{3,5}, which has been implicated in the molecular mechanisms controlling circadian rhythms in mammals⁴⁻⁶. ICER is rhythmically expressed in the pineal gland, with peak levels occurring at night⁵. ICER levels are regulated by light by means of the suprachiasmatic nucleus (SCN); transcription is induced during darkness by adrenergic input to the pineal gland from the SCN, which activates the ICER promoter using cyclic AMP and the transcriptional activator CREB. This induction is transient because ICER represses its own transcription³. Here we show that the response of the CREM gene to adrenergic stimulation is determined by night length. Depending on the photoperiod of the prior entraining cycles, the CREM gene is either subsensitive or supersensitive to induction. This differential responsiveness is controlled by the changing balance between positive (CREB) and negative (ICER) transcriptional regulators. Thus, the transcriptional response of the CREM gene is determined by the memory of past photoperiods.

The cAMP transcriptional response constitutes a key regulatory link between the temporal environment and the neuroendocrine system^{4,7}. Photoperiod regulates neuroendocrine function, enabling mammals to anticipate and adapt appropriately to seasons⁸. We have shown that the CREM gene is strongly induced at night in the pineal gland by clock-derived adrenergic signals^{5,6}. Transcriptional activation of CREM involves binding of phosphorylated CREB to cAMP-responsive elements in the P2 promoter of the gene³. The induced transcript encodes the repressor ICER, which in turn downregulates its own transcription, thus constituting a negative autoregulatory loop³. A subsequent cycle of induction requires a change in the balance between phosphorylation of CREB and the levels of ICER^{3,5,9,10}.

We tested the effect of changing the photoperiod upon ICER expression in the pineal gland. Rats were adapted for 2 weeks in various photoperiods (light-dark (L-D) 3:21; 6:18; 12:12; 18:6; 21:3) (Fig. 1a). We observed a proportional increase with increasing photoperiod in the speed of ICER induction at the beginning of the night (Fig. 1b, c). Under a long photoperiod ICER behaves as an early-response gene, showing a peak of induction by 1.5 h following the onset of the 3-h night. This peak is transient, declining to basal levels within 6 h. Under short photoperiods, the rate of induction is considerably slower: peak levels occurred only 12 and 14 h after the beginning of 16- and 21-h nights. From these results it is evident that the dynamics of ICER expression are versatile because its kinetics are determined by night length.

Is CREM able to adapt to changes in photoperiod by modifying its mode of induction in response to adrenergic stimuli? We have shown that CREM is refractory to cAMP stimulation during the daytime under 12:12 conditions⁵. Rats adapted to long or short photoperiods were challenged with injections of the adrenergic agonist isoprenaline at the beginning of the night (Fig. 2a). CREM is strongly inducible under a long photoperiod (Fig. 2a, lanes 1-4) whereas it is only weakly inducible under a short photoperiod (Fig. 2a, lanes 8-11). We then tested whether CREM inducibility at the end of the night would also adapt to night length (Fig. 2b). Under long photoperiod, CREM is highly induced (Fig. 2b, lower panel, lanes 1 and 2), the gene is refractory to isoprenaline stimulation immediately after a short photoperiod night (Fig. 2b, top panel,

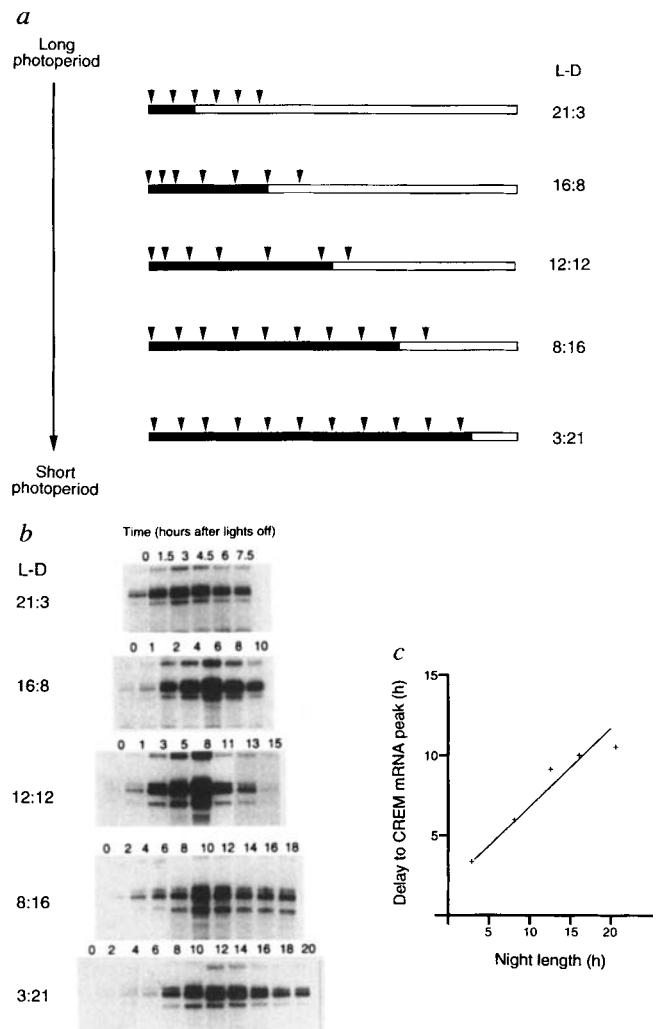


FIG. 1 Changes in CREM expression in different photoperiods. **a**, Representation of different long and short photoperiod regimes: dark bars represent night and light bars daytime. Arrowheads indicate the times at which animals were killed. **b**, RNase protection analysis of ICER mRNA levels during long and short photoperiods using the riboprobe p75 (ref. 5). ICER mRNA levels rise more rapidly under a long photoperiod to reach peak values within 1.5 to 3 h; under a short photoperiod, the equivalent expression of ICER is reached only after 12 to 14 h. Numbers above each lane indicate the time animals were killed (h) after the beginning of each night. **c**, Graph of an apparent linear relation between time required for the CREM gene to reach peak expression and the night length. **METHODS.** RNA extraction from rat pineal glands and RNase protection analysis with CREM-specific probes have been described⁵. Male Wistar rats (40–50 days old; Iffa-Credo, France) were maintained under long and short photoperiods for at least two weeks before analysis³⁰.

lanes 1 and 2). Thus, the CREM transcriptional response may be either supersensitive or subsensitive to the same stimulus. Interestingly, under a short photoperiod, inducibility gradually increases after the night to reach a level after 10 h that is comparable to that of a long photoperiod (Fig. 2b, top panel, lanes 3–12). In contrast, inducibility under long photoperiod remains high (Fig. 2b, lower panel, lanes 3–12). Thus, the duration of the refractory period is determined by the length of the night to which the animals are adapted. It is also conceivable that the refractory period is associated with a reduction in the adrenergic stimulation. Nevertheless, the refractory period can be considered as a memory of night length which constrains the timing of the transcriptional response at the beginning of the following night.

CREM is cAMP-inducible by CREB phosphorylation and is negatively autoregulated by the newly synthesized ICER³. We investigated the role played by these factors during CREM adaptation to changing photoperiods. ICER protein levels show significant differences: the overall levels are considerably higher in

pineal glands from animals adapted to short photoperiods with respect to long photoperiods (Fig. 3a), which is consistent with the slower induction.

We then tested ICER protein levels throughout the day and night during long and a short photoperiods (Fig. 3b, top panel). At

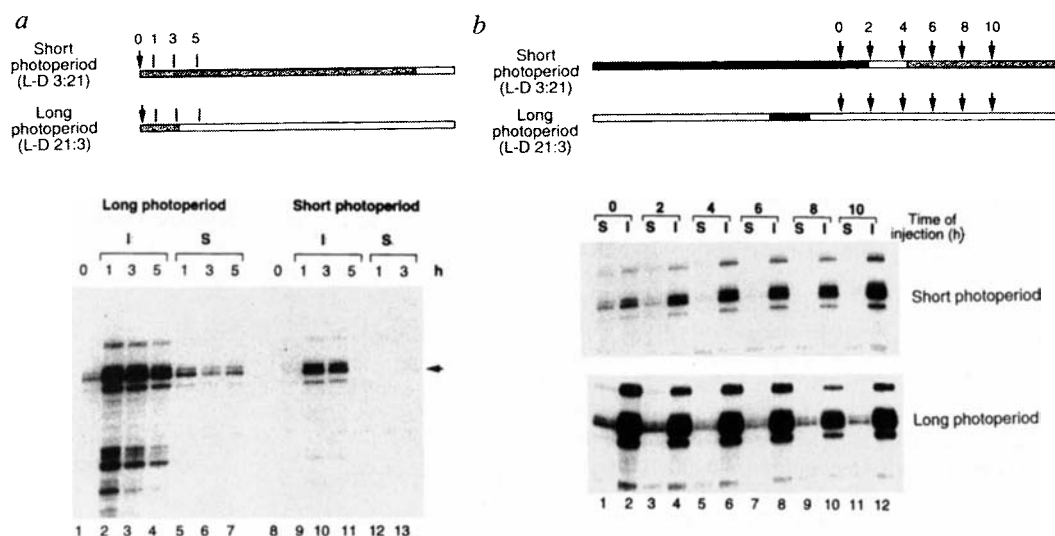
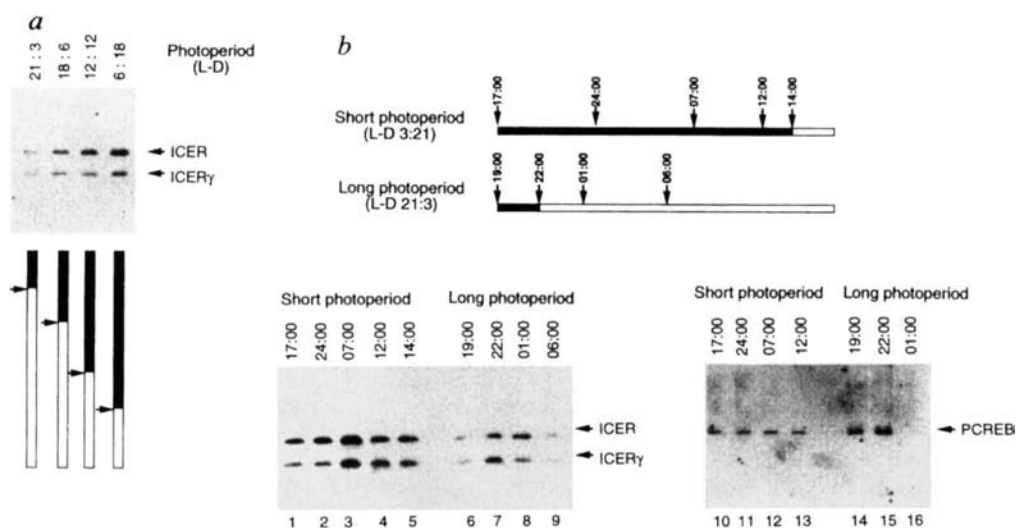


FIG. 2 CREM inducibility is dependent upon photoperiod. *a*, Differential CREM inducibility at the beginning of long or short nights. Animals were maintained under long (L-D, 21:3) and short (L-D; 3:21) photoperiods for two weeks before analysis. During the night of treatment, rats were transferred into constant light to prevent the natural night-time rise in ICER levels⁵. Top panel shows the times of isoprenaline (I) or control saline (S) injections (arrows, and see bottom panel) during the period of the adapted night (grey bars). Vertical lines indicate the times at which injected animals were killed. Lower panel, RNase protection analysis of CREM expression. Induction is maximal between 1 and 3 h during the long photoperiod (lanes 1–4) but is strikingly reduced under short photoperiod conditions (lanes 8–11). *b*, Differential CREM inducibility after long and short nights. Top, animals were adapted to long and short photoperiod conditions and then

injected with isoprenaline (I) or saline (S) after the night at the times indicated (arrows). Animals were killed two hours after injection. In the case of the short photoperiod, the course of injections starts at the end of the night (black bar) and overlaps with the beginning of the next night, at which point the animals were transferred into constant light (grey bar, and see *a*). Bottom, inducibility is maximal at all times tested under long photoperiod (lower, lanes 1–12). In the short photoperiod the gene is not inducible immediately at the end of the night (top, lanes 1 and 2), but gradually gains inducibility at later times (lanes 3–12). (Note that the lower panel has been exposed for longer than the upper.) Equivalent induction is obtained after 10 h for a short photoperiod as in a long photoperiod. Isoprenaline and saline injections, RNA analysis and extractions have all been described⁵.

FIG. 3 Contribution of CREB and ICER proteins to differential inducibility. *a*, Rats were adapted for at least two weeks



to various long and short photoperiods (represented by vertical black and white bars) and then killed at the onset of night (arrows). Pineal gland protein extracts were analysed by western blotting using an anti-ICER antibody (top panel). The ICER and ICER_γ (ref. 3) proteins are indicated; their levels rise with increasing night length. *b*, Top, rats were adapted for at least two weeks to long (L-D 21:3) and short (L-D 3:21) photoperiods and then killed at the times indicated (arrows). Bottom panels: western blots of pineal extracts using anti-ICER antibody (lanes 1–9) and an anti-phosphoCREB antibody (lanes 10–16). ICER and ICER_γ protein bands are indicated that are expressed at higher levels under short than under long photoperiod. Additional analysis confirms that the peak of ICER protein is consistently during the second half of the night. Levels of phosphorylated CREB (PCREB) are low and non-fluctuating under a short photoperiod. Under a long

photoperiod, high levels are visible for the duration of the 3-h night. These then drop to undetectable levels from the onset of the following day. METHODS. Protein extracts and western blots were prepared as described⁵. The anti-ICER antibody has been described^{13,14}. The anti-phosphoCREB antibody¹¹ was obtained from Upstate Biotechnology Inc., New York.

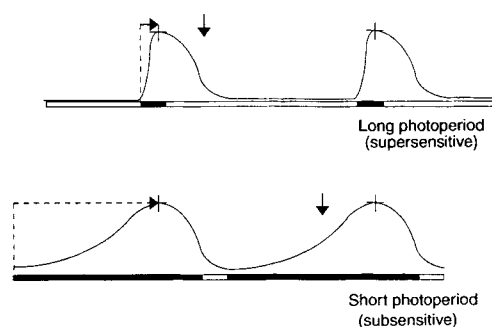


FIG. 4 CREM expression and inducibility following adaptation to short and long photoperiods. Although periodicity of the peak of expression is identical in both photoperiods, the kinetics of the increase are different. Thus, while rhythmicity is conserved, CREM expression increases very rapidly in a long photoperiod, whereas in a short photoperiod the increase is gradual. This is consistent with inducibility studies (Fig. 2). Dashed arrows indicate the time required to reach the peak of expression in different photoperiods; vertical arrows indicate that the timing for maximal inducibility by cAMP following a peak of expression is different depending on the length of the night. Increasing night length dictates a slower gain in inducibility. Therefore the CREM gene has different modes of transcriptional regulation, being supersensitive to cAMP in the long photoperiod and subsensitive during the short photoperiod. Consequently, the length of the refractory period that precedes the subsequent induction depends on the nature of the previous cycle. This adaptive phenomenon represents a 'memory' of transcriptional rhythmicity.

all time points we observed significantly more protein in the pineal glands under short (lanes 1–5) than under long photoperiods (lanes 6–9). Furthermore, in both photoperiods, a peak in protein occurred during the second part of the night which was coincident with the decline in ICER transcript. This peak has a reduced magnitude in short (compare lanes 1 and 3 of Fig. 3b) compared with long photoperiod conditions (compare lanes 6 and 7 of Fig. 3b).

We assayed CREB phosphorylation using an antiphospho-CREB antibody¹¹ (PCREB; Fig. 3b, lanes 10–16). In animals adapted to short photoperiods, low non-fluctuating levels of phosphorylated CREB were detected throughout the whole cycle (Fig. 3b, lanes 10–13). In contrast, during long photoperiods a burst of phosphorylation accompanies the onset of the short night, which persists and then drops to basal levels at the begin-

ning of the next day (Fig. 3b, lanes 14–16). Thus, increasing phosphorylation correlates directly with more rapid CREM induction. A strict relation exists between the changes in the kinetics of CREM expression and the status of the cAMP-responsive activators and repressors.

Transcription factors responsive to cAMP appear to be important in long-term adaptations of neuroendocrine function^{5,12–17}. We have previously shown that CREM is modulated during seasonal changes in spermatogenesis¹⁸. Here we demonstrate that changes in photoperiod have profound effects upon ICER expression in the pineal gland. Differential inducibility of the gene supports a more general notion that the nuclear response to cAMP can be subsensitive or supersensitive. This is in agreement with the original descriptions of differential sensitivity to β -adrenergic stimulation in the pineal gland^{19–24}. Our findings extend this phenomenon to mechanisms regulating gene expression in the nucleus.

CREM rhythmic expression may have the characteristics of an early-response gene (long photoperiod) or have a delayed, dampened response (short photoperiod). We show that the length of the refractory period in CREM inducibility is determined by the photoperiod. Figure 4 is a schematic representation of the different kinetics of CREM expression during long and short photoperiods, displayed in combination with the subsensitive and supersensitive states of inducibility. It should be stressed that the molecular adaptation described here is likely to operate together with the well documented photoperiod-dependent entrainment of the clock system^{25,26}. Furthermore, other photoperiod-dependent physiological changes may contribute to this phenomenon. However, the role of ICER and CREB in the adaptation to a specific state of inducibility is evident from our data (Fig. 3). Furthermore, the involvement of additional nuclear factors in this phenomenon cannot be excluded. However, the levels of CREB-binding protein (CBP; ref. 4) and protein kinase A do not change under different photoperiods (results not shown).

The photoperiod-dependent changes in ICER correlate well with the previously documented seasonal variations in melatonin²⁷. This further strengthens the notion that ICER may be directly involved in melatonin synthesis through the regulation of the serotonin *N*-acetyl transferase gene^{12,28,29}.

Here we have described a remarkable feature of gene expression which enables animals to adapt to temporal changes in their environment. This adaptation involves the establishment and maintenance of different patterns of transcriptional regulation. Thus, the CREB expression adapts to changing day–night rhythms by its experience of previous photoperiods. □

Received 25 October 1995; accepted 28 February 1996.

- Foulkes, N. S., Borrelli, E. & Sassone-Corsi, P. *Cell* **64**, 739–749 (1991).
- Foulkes, N. S., Mellström, B., Benusiglio, E. & Sassone-Corsi, P. *Nature* **355**, 80–84 (1992).
- Molina, C. A., Foulkes, N. S., Lalli, E. & Sassone-Corsi, P. *Cell* **75**, 875–886 (1993).
- Sassone-Corsi, P. *A. Rev. Cell dev. Biol.* **11**, 355–377 (1995).
- Stehle, J. H. et al. *Nature* **365**, 314–320 (1993).
- Stehle, J. H., Foulkes, N. S., Pévet, P. & Sassone-Corsi, P. *Molec. Endocr.* **9**, 706–716 (1995).
- Sugden, D., Vanecek, J., Klein, D. C. & Thomas, T. P. *Nature* **314**, 359–361 (1985).
- Reiter, R. J. *Endocr. Rev.* **12**, 151–180 (1991).
- Sassone-Corsi, P. *Cell* **78**, 361–364 (1994).
- Armstrong, A., Wen, W., Meinkoth, J., Taylor, S. & Montminy, M. *Molec. cell. Biol.* **15**, 1826–1832 (1995).
- Ginty, D. D. et al. *Science* **260**, 238–241 (1993).
- Takahashi, J. *Curr. Biol.* **4**, 165–168 (1994).
- Lalli, E. & Sassone-Corsi, P. *Proc. natn. Acad. Sci. U.S.A.* **92**, 9633–9637 (1995).
- Monaco, L., Foulkes, N. S. & Sassone-Corsi, P. *Proc. natn. Acad. Sci. U.S.A.* **92**, 10673–10677 (1995).
- Yin, J. C. P. et al. *Cell* **79**, 49–58 (1994).
- Huang, Y. Y., Li, X. C. & Kandel, E. R. *Cell* **79**, 69–79 (1994).
- Bourtchuladze, R. et al. *Cell* **79**, 59–68 (1994).
- Foulkes, N. S., Schlöter, F., Pévet, P. & Sassone-Corsi, P. *Nature* **362**, 264–267 (1993).

- Deguchi, T. & Axelrod, J. *Proc. natn. Acad. Sci. U.S.A.* **69**, 2208–2211 (1972).
- Deguchi, T. & Axelrod, J. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2411–2414 (1973).
- Axelrod, J. *Science* **184**, 1341–1348 (1974).
- Romero, J. A., Zatz, M. & Axelrod, J. *Proc. natn. Acad. Sci. U.S.A.* **72**, 2107–2111 (1975).
- Romero, J. A. & Axelrod, J. *Science* **184**, 1091–1092 (1974).
- Romero, J. A. & Axelrod, J. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1661–1665 (1975).
- Illnerova, H. *Pineal Res. Rev.* **6**, 173–217 (1988).
- Illnerova, H. in *Circadian Rhythms in the Mammalian Pineal Gland* 1–105 (Academia, Prague, 1986).
- Illnerova, H. & Vanecek, J. *J. pineal Res.* **2**, 67–78 (1985).
- Coon, S. L. et al. *Science* **270**, 1681–1683 (1995).
- Borjigin, J., Wang, M. M. & Snyder, S. H. *Nature* **378**, 783–785 (1995).
- Sumová, A., Travníková, Z., Peters, R., Schwartz, W. J. & Illnerová, H. *Proc. natn. Acad. Sci. U.S.A.* **92**, 7754–7758 (1995).

ACKNOWLEDGEMENTS. We thank H. Illnerova, M. Rosbash, E. Borrelli, E. Lalli, M. Lamas, L. Monaco, F. Nantel, J. H. Stehle, E. Zazopoulos, K. Tamai, P. Garcia-Villalba, L. Richie and J. S. Lee for discussion and for their support, and E. Heitz and B. Boulay for technical assistance. This study was funded by grants from CNRS, INSERM, CHUR, Rhône-Poulenc Rorer (Bioavenir), and Association pour Recherche sur le Cancer.

CORRESPONDENCE and requests for materials should be addressed to P.S.C. (e-mail paolosc@titus.u-strasb.fr).

A human RNA polymerase II complex associated with SRB and DNA-repair proteins

Edio Maldonado, Ramin Shiekhhattar, Michael Sheldon*, Helen Cho, Ronny Drapkin, Paula Rickert†, Emma Lees†, Carl W. Anderson‡, Stuart Linn§ & Danny Reinberg

Howard Hughes Medical Institute, Department of Biochemistry, Robert Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey, 675 Hoes Lane, Piscataway, New Jersey 08854-5635, USA

† DNAX Research Institute of Molecular and Cellular Biology, Department of Chemistry, Stanford University, Palo Alto, California 94304, USA

‡ Biology Department, Brookhaven National Laboratory, Upton, New York 11973, USA

§ Department of Biochemistry, University of California at Berkeley, California 94720, USA

We report here the isolation of a human RNA polymerase II complex containing a subset of the basal transcription factors and the human homologues of the yeast SRB (for suppressors of RNA polymerase B) proteins¹⁻³. The complex contains transcriptional coactivators and increases the activation of transcription. In addition, some components of the RNA polymerase II complex participate in DNA repair.

A complex containing RNA polymerase II (RNAPII) was purified from HeLa cell nuclear extract using a combination of conventional and affinity chromatography (Fig. 1). Analysis of fractions derived from the last step of purification (gel filtration on a Sepharose CL-4B column) by SDS-PAGE followed by silver staining demonstrated that RNAPII is a component of a large complex with a relative molecular mass of ~2,000K (Fig. 1a). To determine the size of the purified RNAPII complex more accurately, we compared the sedimentation of the RNAPII complex to that of 60S and 80S ribosomes in a sucrose gradient. Core RNAPII was used as a control, sedimenting at the top of the gradient (Fig. 1b). In contrast, the RNAPII complex sedimented faster, close to the position of the 60S ribosome subunit (Fig. 1b). Most of the RNAPII in nuclear extracts sedimented in a broad peak, spanning from the RNAPII complex to heavier forms (Fig. 1b).

Quantitative western blot analysis scoring for the presence of the different general transcription factors (GTFs) in the purified RNAPII complex (Fig. 1a, and data not shown) demonstrated that transcription factor (TF) IIE (18 fmol μl^{-1}) and TFIIF (14 fmol μl^{-1}) were stoichiometric with subunits of RNAPII (18 fmol μl^{-1}), but that TFIIF (2 fmol μl^{-1}) was limiting (data not shown). TATA-binding protein (TBP), TFIIB and TBP-associated factors were absent from the complex (Fig. 2, and data not shown). The complex also contained stoichiometric amounts of YY1 (13 fmol μl^{-1}) (Fig. 1a), a sequence-specific DNA-binding protein implicated in initiator-dependent transcription from TATA-less promoters⁴.

We investigated whether the human homologues of yeast SRB proteins are associated with the RNAPII complex. There is significant homology between yeast SRB10/SRB11 and human cyclin C/CDK8 (ref. 5) and a recent report has described the existence of a mammalian RNAPII holoenzyme containing SRB7 (ref. 6). We isolated a complementary DNA clone encoding the human homologue of yeast SRB7 and used the recombinant polypeptide to generate antibodies (data not shown). Western blot analysis of the conventionally (Fig. 1a) and affinity-purified (see below) RNAPII complex revealed the presence of human SRB7 and cyclin C/CDK8.

The complex was immunopurified using monoclonal antibodies directed against the large subunit of TFIIF (RAP74). SDS-PAGE analysis of the column eluate revealed the presence of core RNAPII subunits and a number of other polypeptides (Fig. 1c). Fractionation of this eluate by gel filtration resulted in a similar elution profile to that of the conventionally purified complex (data not shown), indicating that the affinity-purified eluate is also a large large complex. Western blot analysis of the affinity-purified RNAPII complex confirmed the presence of RNAPII, TFIIE, TFIIF, cyclin C, CDK8 and YY1 (Fig. 1d). These proteins were absent from the eluate of the control column (Fig. 1d).

The functional analysis supports the above results and demonstrates that the RNAPII complex, but not the core RNAPII, can direct transcription in the absence of TFIIE and TFIIF (Fig. 2a). TBP and TFIIB are required for transcription regardless of the form of RNAPII used (Fig. 2b). To analyse whether a functional form of TFIIF was present in the complex, reactions were reconstituted in the absence of TFIIF, TFIIE and TFIIF. There was transcription in the absence of added TFIIF which was inhibited by antibodies against the ERCC3 subunit of TFIIF, but not by preimmune sera (Fig. 2c). This inhibition was specific as it was neutralized by excess TFIIF (lane 4), but not TBP or TFIIB (lane 3). Addition of TFIIF stimulated transcription, indicating that TFIIF was functionally limiting in the RNAPII complex (lane 6).

Next we studied the RNAPII complex in activated transcription *in vitro*. Activated levels of transcription using core RNAPII and the acidic activator GAL4-VP16 require cofactors such as PC4 (refs 7, 8) and cofactor A (Cof-A), a new activity (Fig. 3a, lanes 1-3, and unpublished results). In contrast to core RNAPII, the RNAPII complex did not require the addition of cofactors to achieve activated levels of transcription, indicating that the complex contains factors capable of mediating transcriptional activation (Fig. 3a, lanes 4-6). Western blot analysis of the RNAPII complex indicated the absence of PC4. However, a coactivator activity has been reported to be associated with the C-terminal domain (CTD) of a yeast RNAPII holoenzyme^{1,2}. To analyse whether the coactivator activity of the human RNAPII complex is associated (directly or indirectly) with the CTD, the complex was fractionated on an anti-CTD antibody column. Antibodies against the CTD dissociate protein-CTD interactions. Functional analysis combined with western blots revealed that the flow-through fraction did not contain RNAPII, yet could mediate transcriptional activation when added to core RNAPII (Fig. 3b). Together, these experiments indicate that the RNAPII complex supports activated transcription and that a coactivator activity is associated with the CTD of RNAPII.

It was previously reported that TFIIF has dual roles in transcription and DNA nucleotide excision repair (NER)⁹⁻¹¹. Because the RNAPII complex contains TFIIF, we analysed the RNAPII complex in a NER assay. The RNAPII complex only weakly complemented a xeroderma pigmentosum D (XPD) cell-free extract, consistent with limiting amounts of TFIIF in the complex (see above). Surprisingly, the RNAPII complex could efficiently complement XPF-, XPG- and XPC-repair-deficient extracts (Fig. 4a). Furthermore, western blot analysis of the RNAPII complex using anti-XPG, -XPC and -ERCC1 (XPF component) antibodies revealed their presence in the RNAPII complex (data not shown). As the yeast homologues of ERCC1/XPF (RAD1/RAD10) have a dual role in nucleotide excision repair and recombinational/double-strand break repair¹²⁻¹⁴, we examined the RNAPII complex for other components involved in DNA double-strand break and recombinational repair. The Sepharose CL-4B column (Fig. 4b) revealed the coelution of the Ku regulatory subunits of DNA-PK, a protein kinase critical for DNA double-strand break repair and V(D)J recombination¹⁵⁻¹⁹, and human RAD51 (refs 20-22) with the largest subunit of RNAPII (RPB1) and with RNAPII complex transcriptional activity (for transcription, see Fig. 1a). Ku interacts with DNA polymerase ϵ

* Present address: Department of Developmental Neurobiology, St Jude Children's Research Hospital, 312 North Lauderdale, Memphis, Tennessee 38105, USA.

(Pol ϵ)²³, and Pol ϵ participates in the process of DNA-repair synthesis^{24,25}. We therefore examined the RNAPII complex for DNA polymerase activity and immunoreactivity towards different components of DNA-repair synthesis. Western blot analysis of the Sepharose CL-4B fractions revealed the presence of DNA Pol ϵ ,

replication protein A (RPA) and replication factor C (RFC). The RNAPII complex was also active in an assay measuring DNA polymerase ϵ and RFC activity (data not shown). Absent from the complex were ERCC6, DNA pol δ , topoisomerase I. Quantitative western blot analysis revealed that HRAD51, RPA and Ku are

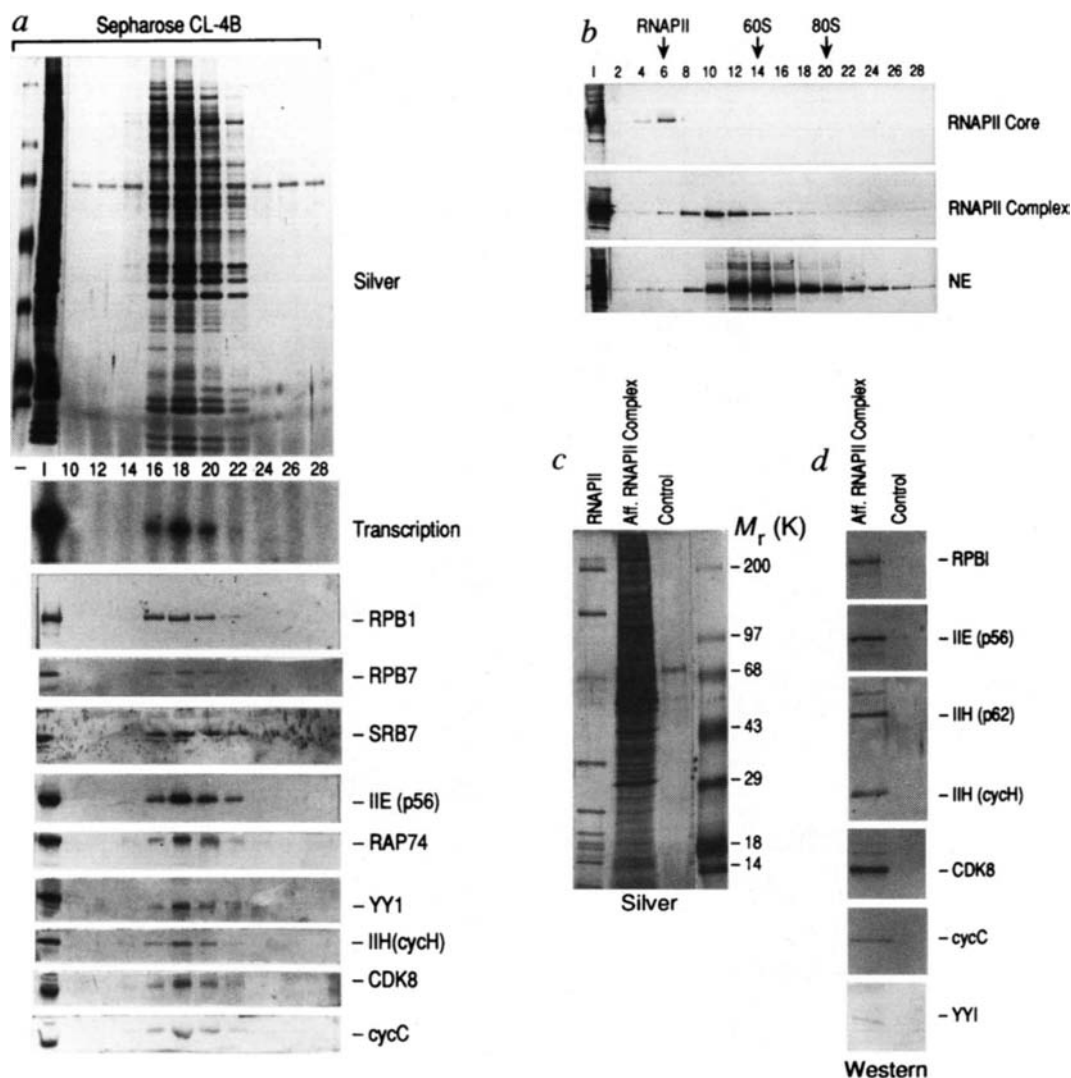


FIG. 1 Purification and characterization of the RNAPII complex. *a*, Analysis of the RNAPII complex purified by conventional chromatography from the last gel filtration step (Sepharose CL-4B). The complex was purified after an activity capable of directing specific transcription in an assay reconstituted with recombinant TBP, recombinant TFIIB and native TFIIF. SDS-PAGE followed by silver staining and western blot analysis of the column fractions using the antibodies against factors indicated on the right M_r markers to the far left of the silver staining are the same as in *c*. *b*, Input to the column (Mono S, 5 μ l, 1 μ g protein). Numbers, Fractions of the CL-4B column (10 μ l). Transcription assays were as described in Fig. 2a, *b*. Sucrose gradient sedimentation analysis of the RNAPII complex. The sedimentation of the RNAPII complex and RNAPII in HeLa cell nuclear extracts was compared to that of core RNAPII, rat 80S and 60S ribosome. *c*, Silver staining of an SDS-PAGE gel containing the anti-RAP74 affinity-purified RNAPII complex. From left to right are core RNAPII (2 μ l, 0.2 μ g); Alkyl-superoxide fraction) purified from HeLa nuclear pellets²⁸, the eluates from the anti-RAP74 monoclonal antibodies (100 μ l eluate, representing $\sim$ 0.5 μ g protein) and the anti- β -galactosidase control column. *d*, Western blot analysis of the affinity-purified RNAPII complex (100 μ l eluate) using antibodies against factors shown to the right.

METHODS. A full description of the conventional and affinity purification schemes as well as sucrose gradient sedimentation of the RNAPII complex

is contained in Supplementary Information. RNAPII complex was purified from HeLa nuclear extract. An aliquot of the Mono-S fraction (0.6 ml, 0.12 mg) was concentrated against 50% glycerol (to 200 μ l) in buffer H (20 mM HEPES-KOH, pH 7.9, 0.5 mM EDTA, 5 mM DTT, 10% glycerol, 0.01% NP-40, 0.1 mM PMSF, 1 μ g ml⁻¹ pepstatin A), containing 0.5 M KCl and was fractionated on a gel filtration Sepharose CL-4B column (12 ml column, 0.7 $\times$ 30 cm; Pharmacia). The M_r fractionation of the column ranges between 60K and 20,000K for globular proteins. First, 2 ml was collected in a single fraction, and then fractions of 0.2 ml were collected. The void volume of the column was determined by using polystyrene (M_r 30,000K, Polysciences Inc.) which peaks at fraction 9. The elution of RNAPII complex was compared to that of the M_r markers (blue dextran, M_r 2 $\times$ 10⁶; thyroglobulin, M_r 0.7 $\times$ 10⁶; alcohol dehydrogenase, M_r 0.15 $\times$ 10⁶). The peak of RNAPII complex activity elutes with an apparent M_r of 2,000K (the peak of the blue dextran marker is at fraction 18). For silver staining, 15 μ l of each fraction was loaded on a SDS-PAGE. For comparison, 15 μ l (3 μ g) of the input was loaded on the same gel. Transcription assays were performed with 10 μ l of each fraction. For the western blots, 10 μ l of the input and 30 μ l of each fraction was analysed and the proteins were detected with the indicated rabbit polyclonal antibodies and visualized using alkaline-phosphatase-conjugated secondary antibodies (Promega).

FIG. 2 Transcription factors IIE, IIF and IIH are present in the RNAPII complex. *a*, Transcription assays were reconstituted on the Ad-MLP promoter with recombinant human TFIIB (20 ng), recombinant human TBP (10 ng) and purified TFIIF (500 ng, phenyl-Superose fraction)¹⁰. Equal western blot units (1, 2, 4) of RNAPII complex (lanes 1–3) and core RNAPII (lanes 4–6) were analysed. Purified TFIIF (60 ng, phenyl-superose fraction)²⁹ was added to the core RNAPII (lanes 7–9). A partially purified fraction (500 ng, Superdex 200) of TFIIE/TFIIF was added to the core RNAPII in lane 10. Western blot units of RNAPII were determined using anti-CTD antibodies. One unit was arbitrarily defined as the amount of RNAPII complex which gives a signal equivalent to 0.04 pmol of the largest subunit of purified core RNAPII (hRPB1). *b*, Transcription by the RNAPII complex requires TFIIB and TBP. Transcription assays were reconstituted with 5 units (5U) of the RNAPII complex or as indicated. TFIIB (lane 1), TBP (lane 2) and RNAPII complex (lane 3) were omitted. The indicated units of RNAPII complex were added in the presence of TFIIB, TBP and TFIIF in lanes 4–6. *c*, Transcription directed by the RNAPII complex is sensitive to anti-ERCC3 antibodies. Reactions are described above. The ERCC3 polypeptide was not detected by western blot in the RNAPII complex preparations, most probably because of its low abundance. Transcription reactions were reconstituted as indicated. In lanes 2–4, the RNAPII complex (Mono-S fraction, 1.2 μ g) was incubated for 30 min at 4 °C with 0.6 μ g affinity-purified anti-ERCC3 polyclonal antibodies. Lane 5, the complex was treated with preimmune IgG (1 μ g) as in lane 2. In lane 3 the amount of TBP/TFIIB was increased (40 ng each, this amount does not inhibit transcription). Lanes 4 and 6 were supplemented with TFIIF (500 ng, phenyl-superose fraction).

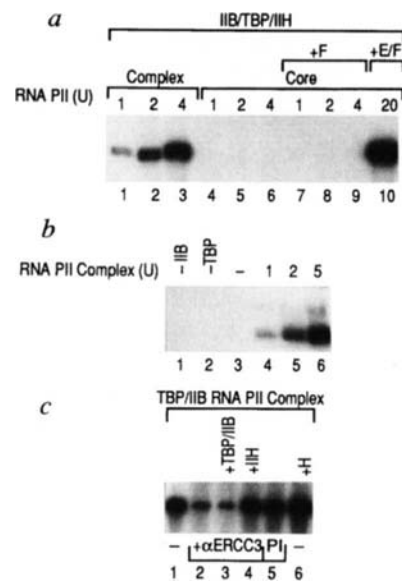


FIG. 3 The RNAPII complex contains activities involved in transcriptional activation. *a*, The RNAPII complex does not require additional cofactors (other than TAFs) to respond to the activator GAL4-VP16 (Act). Transcription reactions were reconstituted using two different templates containing the Ad-MLP (20 ng of each). Plus, Transcript generated by pG5MLT, a construct based on the adenovirus MLP with 5 GAL4 sites upstream of the TATA box. Minus, Transcript generated by p210, a similar construct lacking GAL4 sites driving a shorter G-less cassette. Transcription reactions were reconstituted with: affinity-purified eTFIID (50 ng), TFIIF/H (500 ng, phenyl-superose fraction), core RNAPII (DEAE-5PW fraction, 150 ng)²⁸, recombinant TFIIA (40 ng), recombinant TFIIB (20 ng), and recombinant TFIIE (20 ng). Reactions also contained recombinant PC4 (10 ng)^{7,8} and Cof-A (1 μ l of HQ50 fraction), collectively referred to as Cofs, and GAL4-VP16 (10 ng) as indicated. In lanes 4–6, RNAPII complex (Mono-S fraction, 600 ng) was used. Additions and omissions are as shown. *b*, The RNAPII complex contains a CTD-associated coactivator activity. Reactions were reconstituted as described in *a*. The input (I), RNAPII complex Mono-S fraction, 600 ng) and flow-through (Fth, 2 μ l; 400 ng) of an anti CTD column were assayed for their activity as cofactors when core RNAPII was used (lanes 1–4).

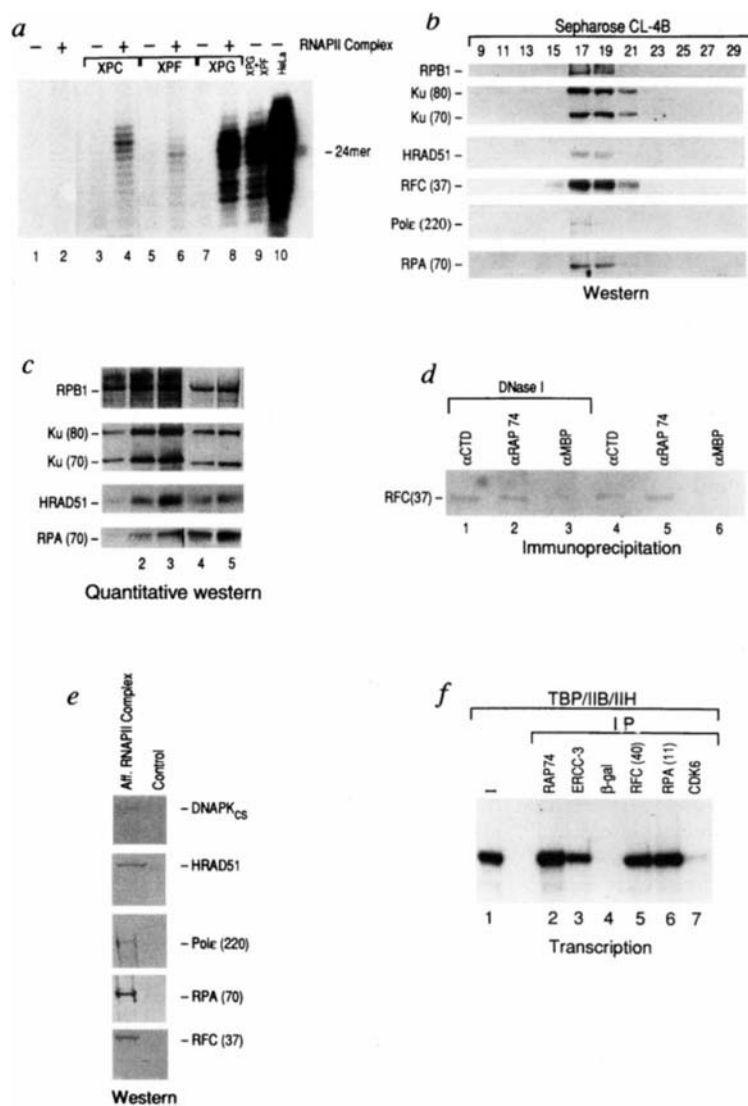
METHODS. The coactivator fraction was separated from the RNAPII complex by anti-CTD affinity chromatography. The RNAPII complex from the MonoS fraction (600 μ l, 120 μ g) was dialysed against buffer H containing 0.5 M KCl and incubated for 45 min at 4 °C with anti-CTD antibodies (8WG16, Promega) bound to protein A-agarose (150 μ l, 150 μ g antibodies). The flowthrough from the affinity column was recovered, and the column was washed with the above buffer. The flow-through and the wash were pooled and concentrated to 500 μ l by dialysing against buffer H containing 0.1 M KCl and 50% glycerol.

present in stoichiometric amounts (20 fmol μ l⁻¹ for each protein) with the largest subunit of RNAPII (Fig. 4c).

The association of DNA-repair factors with the RNAPII complex was further analysed by immunoprecipitation and immunoaffinity purification. Western blot analysis of the affinity-purified RNAPII complex revealed the presence of DNA-PKcs (catalytic subunit), HRAD51, DNA polymerase ϵ , RPA70K subunit, and RFC37K subunit (Fig. 4e). These proteins were not detected in the control β -galactosidase column eluate (Fig. 4e). Furthermore, antibodies against RNAPII or TFIIF, but not control antibodies, immunoprecipitated RFC (Fig. 4d) and other polypeptides of the RNAPII complex (data not shown). The integrity of the RNAPII complex was not affected by DNase I treatment (Fig. 4d; or micrococcal nuclease, or if the complex was fractionated on a

gel-filtration column containing ethidium bromide; data not shown), thereby ruling out the possibility that the complex is held together by nucleic acids. The immunoprecipitated complexes were also analysed functionally in a transcription assay devoid of RNAPII, TFIIE and TFIIF. Antibodies to RFC, RPA, TFIIF (RAP74) and TFIIF (ERCC3) immunoprecipitated a transcriptionally active RNAPII complex, whereas immunoprecipitates with control antibodies had no RNAPII activity (Fig. 4f). Antibodies against CDK7, a subunit of TFIIF, have previously been used to isolate a putative RNAPII holoenzyme containing all the GTFs²⁶. TFIIF is known to interact independently with RNAPII, TFIID, TFIIB, TFIIE, TFIIF and other factors²⁷, therefore immunoprecipitation of all the GTFs and RNAPII by anti-CDK7 antibodies does not unequivocally

FIG. 4 RNAPII complex contains components required for DNA repair. **a**, Nucleotide excision assay scoring for complementation activity of RNAPII complex (3–4 μ l; 0.6–0.8 μ g, MonoS) using XPC- (lanes 3 and 4), XPF- (lanes 5 and 6), and XPG- (lanes 7 and 8) deficient extracts. Lanes 1 and 2 are in the absence of any extract, lane 9 is a mixture of XPG and XPF extracts and lane 10 is the control HeLa extract. Excision assays were described previously³⁰. **b**, Western blot analysis of Sepharose CL-4B gel filtration column fractions (50 μ l, see Fig. 1a) using antibodies to proteins delineated on the left. **c**, Quantitative western blot analysis of RNAPII complex (lanes 4, 5, 10 μ l (2 μ g) and 20 μ l (4 μ g) of MonoS pool, respectively). Lanes 1, 2 and 3 are: RNAPII, 0.25, 0.5, 0.75 pmol; Ku, 0.14, 0.7, 1.4 pmol; HRAD 51, 0.06, 0.3, 0.6 pmol; RPA, 0.05, 0.25, 0.5 pmol, respectively. **d**, Immunoprecipitation using RNAPII (lanes 1 and 4) and RAP 74 (lanes 2 and 5) antibodies revealed the presence of the 37K subunit of RFC. Lanes 3 and 6 are the immunoprecipitates of the control MBP antibody. Lanes 1–3 denote immunoprecipitation in the presence of 100 units of DNase I as described³¹. The Aca22 (8 μ g) fraction was used as an input for the immunoprecipitation experiments as described¹⁰. Quantification of western blots is described in Supplementary Information. **e**, Analysis of components of DNA repair in the affinity-purified RNAPII complex purified by monoclonal anti-RAP74 antibodies. The same quantity of protein as in Fig. 1c was used. **f**, Antibodies against components of the RNAPII complex can immunoprecipitate a complex which requires the addition of TBP, IIB and IIF for specific basal transcription. Lane 1 represents transcription of 15% of the input fraction. Lanes 2–7 contain transcription from the immunoprecipitates of the corresponding antibodies (monoclonal antibodies against RAP74, ERCC-3 and β -galactosidase (Promega), and polyclonal antibodies against RFC (40), RPA (11) and CDK6 (Santa Cruz Biotechnology) as indicated. To immunoprecipitate the RNAPII complex, 3 μ g of different antibodies (as indicated on the top) were bound to 15 μ l of protein G-agarose beads for 90 min, then extensively washed with buffer H containing 0.1% NP40, 1 mM DTT, 100 mM KCl and 20% glycerol. Aca22-purified RNAPII complex fraction (0.2 ml, 40 μ g) was incubated with each antibody at 4 °C for 2 h and extensively washed with buffer H as described. Transcription from the AdMLP driving a G-less cassette was reconstituted by the addition of recombinant TBP (10 ng), recombinant IIB (20 ng), TFIIF and 15 μ l of each immunoprecipitate. Anti- β -galactosidase and anti-CDK6 immunoprecipitates were assayed as negative controls.



demonstrate the presence of a large RNAPII complex and may represent individual subcomplexes to transcription factors and TFIIF. The results presented here clearly demonstrate the existence of such an RNAPII complex which contains both transcriptional coactivators and components with roles in DNA repair. □

Received 14 September 1995; accepted 11 March 1996.

- Koleske, A. J. & Young, R. A. *Nature* **368**, 466–469 (1994).
- Kim, Y.-J., Bjorklund, S., Li, Y., Sayer, M. H. & Kornberg, R. D. *Cell* **77**, 599–608 (1994).
- Thompson, G. M. & Young, R. A. *Proc. natn. Acad. Sci. U.S.A.* **92**, 4587–4590 (1995).
- Seto, E., Shi, Y. & Shenk, T. *Nature* **354**, 241–245 (1991).
- Tassan, J.-P., Jaquenoud, M., Leopold, P., Schultz, S. J. & Nigg, E. A. *Proc. natn. Acad. Sci. U.S.A.* **92**, 8871–8875 (1995).
- Chao, D. M. *et al.* *Nature* **380**, 82–85 (1996).
- Ge, H. & Roeder, R. G. *Cell* **78**, 513–523 (1994).
- Kretschmar, M., Kaiser, K., Lottspeich, F. & Meisterernst, M. *Cell* **78**, 525–534 (1994).
- Wang, Z. *et al.* *Nature* **368**, 74–75 (1994).
- Drapkin, R. *et al.* *Nature* **368**, 769–772 (1994).
- Van Vuuren, A. J. *et al.* *EMBO J.* **13**, 1645–1653 (1994).
- Ivanov, E. L. & Haber, J. E. *Molec. cell. Biol.* **15**, 2245–2251 (1995).
- Habraken, Y., Sung, P., Prakash, L. & Prakash, S. *Nature* **371**, 531–534 (1994).
- Bardwell, A. J., Bardwell, L., Tomkinson, A. E. & Friedberg, E. C. *Science* **265**, 2082–2085 (1994).
- Smider, V., Rathmell, W. K., Lieber, M. & Chu, G. *Science* **266**, 288–291 (1994).
- Taccioli, G. E. *et al.* *Science* **265**, 1442–1445 (1994).
- Blunt, T. *et al.* *Cell* **80**, 813–823 (1995).
- Kirchgesner, U. C. *et al.* *Science* **267**, 1178–1183 (1995).
- Less-Miller, S. P. *et al.* *Science* **267**, 1183–1185 (1995).
- Adzuma, K., Ogawa, T. & Ogawa, H. *Molec. cell. Biol.* **4**, 2735–2744 (1984).

- Shinohara, A., Ogawa, H. & Ogawa, T. *Cell* **69**, 457–470 (1992).
- Aboussekhra, A., Chanet, R., Adjiri, A. & Fabre, F. *Molec. cell. Biol.* **12**, 3224–3234 (1992).
- Hwang, P.-I., Chui, G., Asahara, H. & Linn, S. *FASEB J.* **9**, A1398 (1995).
- Nishida, C. P., Reinhard, P. & Linn, S. *J. biol. Chem.* **263**, 501–510 (1988).
- Aboussekhra, A. *et al.* *Cell* **80**, 859–869 (1995).
- Ossipow *et al.* *Cell* **83**, 137–146 (1995).
- Gerard, M. *et al.* *J. biol. Chem.* **266**, 20940–20945 (1991).
- Lu, H., Flores, O., Weinmann, R. & Reinberg, D. *Proc. natn. Acad. Sci. U.S.A.* **88**, 10004–10008 (1991).
- Flores, O., Maldonado, E. & Reinberg, D. *J. biol. Chem.* **264**, 8913–8921 (1989).
- Huang, J. C., Svoboda, D. L., Reardon, J. T. & Sancar, A. *Proc. natn. Acad. Sci. U.S.A.* **89**, 3664–3668 (1992).
- Lai, J.-S. & Herr, W. *Proc. natn. Acad. Sci. U.S.A.* **89**, 6958–6962 (1992).

SUPPLEMENTARY INFORMATION. Requests should be addressed to Mary Sheehan at the London editorial office of *Nature*. Supplementary Information is also available at the *Nature* Web site <http://WWW.nature.com>.

ACKNOWLEDGEMENTS. R.S. and E.M. contributed equally to this work. We thank T. Carter, S. Jackson and J. Hurwitz for antibodies; E. Golemis for anti-RPB7 antibodies; H. Flores-Rozas for anti-RFC antibodies; Z. O. Pan for anti-Pol δ antibodies; E. Gibbs for anti-RPA antibodies; T. Matsunaga for anti-XPG antibodies; J. Reardon for anti-XPC antibodies; D. Mu for anti-ERCC1 and PCNA antibodies; R. Young and D. Chao for anti-yeast SRB antibodies and stimulating discussions; Y. Li for DNA polymerase α assays; L. Weis for anti-YY1 antibodies; M. Jaquenoud and E. Nigg for CDK8 antibodies; E. I. Golub, G. Reddy and C. M. Radding for anti-RAD51 antibodies; Austral Biological-Bios, Chile for RAP74 antibodies; U. Maitra for ribosome subunits, C. Selby for DNA-repair assays; and C. Selby, A. Sancar, J. Hurwitz, G. Orphanides and D. Ward for critical reading of the manuscript. The IMAGE Consortium provided the human EST (Homo sapiens cDNA clone 33040'5). R.S. and M.S. are supported by NIH postdoctoral grants, R.D. by an NIH predoctoral grant. This work was supported by grants from the NIH and the Howard Hughes Medical Institute to D.R.

CORRESPONDENCE and requests for materials should be addressed to D.R. (e-mail: Reinberg@mbcl.rutgers.edu). The cDNA sequence of the human homologue of SRB7 is lodged with EMBL, accession number U52960.

Structure of *Bordetella pertussis* virulence factor P.69 pertactin

Paul Emsley, Ian G. Charles*, Neil F. Fairweather†, Neil W. Isaacs

Department of Chemistry, University of Glasgow, Glasgow G12 8QQ, UK
 * Cruciform Project, University College London, 140 Tottenham Court Road, London W1P 9LN, UK

† Department of Cell Biology, Glaxo-Wellcome Research Laboratories, Beckenham, Kent, BR3 3BS, UK

A NEW generation of whooping-cough vaccines contain P.69 pertactin, a surface-exposed domain of an outer membrane protein expressed by the virulent bacterium *Bordetella pertussis*¹⁻³. This protein is a virulence factor that mediates adhesion to target mammalian cells, a reaction that is in part mediated by an RGD sequence. The X-ray crystal structure of P.69 pertactin has been determined to 2.5 Å. The protein fold consists of a 16-stranded parallel β -helix with a V-shaped cross-section, and is the largest β -helix known to date. Several between-strand weakly conserved amino-acid repeats form internal and external ladders. The structure appears as a helix from which several loops protrude, which contain sequence motifs associated with the biological activity of the protein. One particular (GGXXP)₅ sequence is located directly after the RGD motif, and may mediate interaction with epithelial cells. The carboxy-terminal region of P.69 pertactin incorporates a (PQP)₅ motif loop containing the major immunoprotective epitope.

P.69 pertactin (known previously as P.69 owing to its apparent relative molecular mass (M_r) of 69K on an SDS gel) is the 60.4K amino-terminal external domain of a 93K precursor polypeptide (known as P.93 pertactin) encoded by the *B. pertussis* gene *pm*⁴. The P.69 pertactin domain can be released from the cell surface by acid⁵ or heat treatment⁶, suggesting cleavage of a peptide bond in non-physiological conditions. The 30K carboxy-terminal domain is present within the outer membrane, and is required for the correct localization of the P.69 pertactin domain on the cell surface⁶.

Pertactin has structural features in common with other molecules that form protein-protein interactions and are involved in cell binding. One such feature is an Arg-Gly-Asp (RGD) tripeptide motif, which is the cell-attachment site of various mammalian adhesion proteins such as fibronectin, vitronectin, fibrinogen and von Willebrand factor⁷. Filamentous haemagglutinin is known to bind to leukocyte integrin complement receptor 3 (CR3)⁸ by

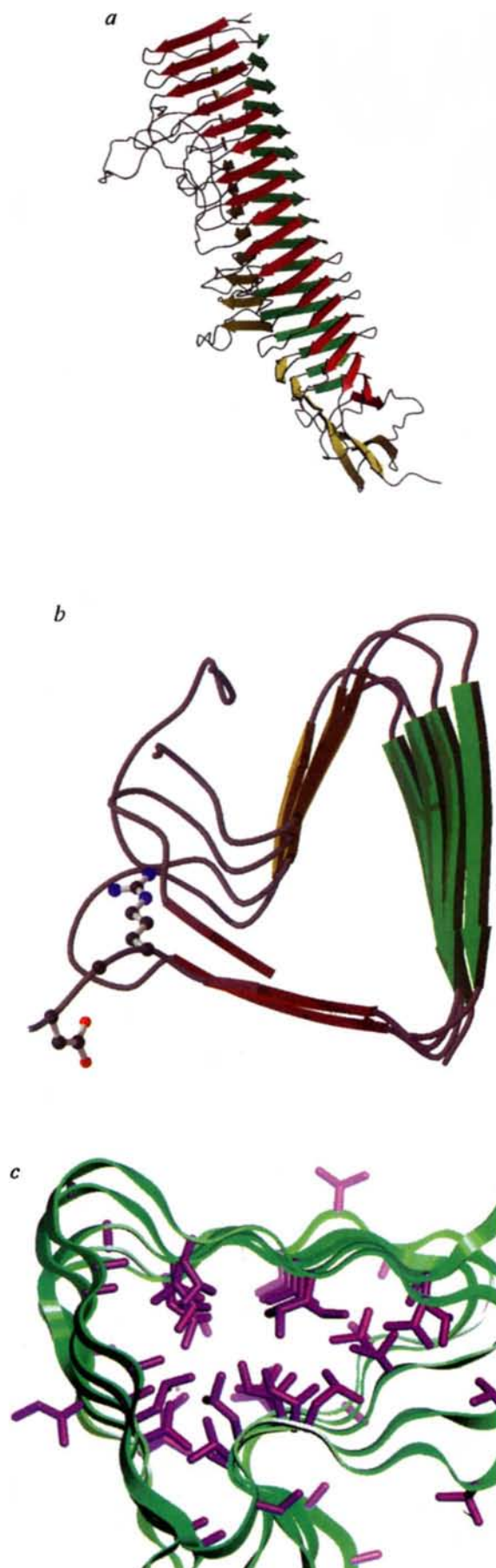


FIG. 1 *a*, Topology of P.69 pertactin. Using the same sheet nomenclature as ref. 12, Sheet 1 is green, Sheet 2 is red, and Sheet 3 is yellow. As with other β -helix proteins^{14,27}, the β -strands have only a shallow right-handed twist. *b*, The RGD site of P.69 pertactin occurs approximately halfway down the length of the helix, and is at the N terminus of a loop. The electron density associated with the immediately following PRRs is poor. It is possible that some of the prolines could have the *cis* conformation, but all were modelled as *trans*-prolines. *a* and *b* were prepared using MOLSCRIPT²⁸. *c*, Internal sidechain stacking. A sample of the aliphatic residue stacking, with aliphatic residues coloured purple, and the main chain represented as a green ribbon. The internal packing is due largely to aliphatic residues, most of which are packed internally. *c* was produced using MolViewer (M. J. Hartshorn, private communication).

means of an RGD sequence. The RGD sequence of P.69 pertactin mediates adhesion to CHO cells⁹, suggesting that P.69 pertactin can also bind integrin.

Proline-rich regions (PRRs) occur widely in proteins with binding activity. It has been suggested that they provide fast, weak, non-stoichiometric but functionally important binding sites¹⁰. P.69 pertactin has two PRRs consisting of tandem repeats: one, following an RGD sequence, consists of the sequence (GGXXP)₅; the other, C-terminal PRR, contains the motif (PQP)₅, and is known to contain the immunodominant region of the molecule¹¹.

P.69 pertactin is a monomer folded into a single domain which is almost entirely three-stranded β -helix¹² (Table 1; Fig. 1a; Fig. 2). Towards the C terminus, the structure forms two rungs of two-stranded β -roll. Unlike that previously seen in alkaline protease¹³, the turns do not incorporate calcium-binding sites, and are rather more extended and variable in structure.

Several loops protrude from the helix. One such loop contains the RGD tripeptide followed in the sequence by a PRR (Fig. 1b). Both of these structural features occur in proteins exhibiting binding activity¹⁰ and are consistent with the observation that P.69 pertactin functions as an adhesin. The 33 residues of the (GGXXP)₅ PRR forms a loop that extends towards the N terminus before returning to the β -helix structure.

There was little indication *a priori* of repetitive structure (other than the PRRs) from the sequence. The interior of the helix comprises largely aliphatic hydrophobic residues, unlike pectate

lyase (PelC), the first β -helix protein to be discovered, which has many Tyr and Phe residues¹⁴. In P.69 pertactin, there are several instances of these aliphatic hydrophobic residues forming stacks (for example, Val, Leu, Ile and Ala ladders) (Fig. 1c), but at the interface between the three-stranded and two-stranded rungs, internal aromatic stacking, comprised of two Trps, does occur. On the exterior of the helix, there is a stack of 12 residues, 6 of which are Thr.

The long linear form of pertactin protruding from the cell surface of the bacterium may be well suited to promoting cell-cell adhesion because it has the potential to present a large surface area to target mammalian cells. It may therefore promote binding through many relatively weak nonspecific interactions, in addition to the RGD motif.

The structure has implications for other PRR and leucine-rich repeat (LRR) proteins, such as the cell-adhesion molecule glycoprotein 1b- α and other members of the integrin family⁷. The outer membrane proteins from *B. pertussis*, BrkA (a serum-resistance protein) and tracheal colonization factor (TcfA), have considerable amino-acid similarity with P.69 pertactin over their C-terminal domains, and they also contain RGD sequences. BrkA also shows significant N-terminal homology, but lacks PRRs; TcfA shows little homology at the N terminus, but is highly proline rich (20%). It is likely that BrkA has a β -helical structure. This motif could be important in outer membrane proteins not only for their translocation across the membrane, but also for their roles as virulence factors. Analysis of the structure of P.69 pertactin will

TABLE 1 Crystallographic data

(a) Crystallographic statistics								
Data set	Resolution (Å)	Completeness (%)	Average multiplicity	R_{merge}^* (%)	Sites	Phasing power†	R_{cull} (Anom. R_{cull})‡	
Native	2.5	99.4	4.8	6.8				
PtCl ₄	3.1	99.8	4.4	8.5	3	1.1	0.74(0.7)	
Pt-terpyridine	3.5	99.4	8.3	3.7	3	0.4	0.87(0.9)	
(b) Merging and model R -factors								
Resolution	4.8	4.0	3.5	3.2	2.95	2.8	2.66	2.54
R_{merge}	3.2	3.5	4.5	6.3	10.0	14.4	21.3	24.4
R_{crist}	19.4	18.1	19.6	23.3	23.7	25.1	26.9	28.9
R_{free}	27.8	25.0	23.8	30.8	31.8	36.8	37.7	32.2

P.69 pertactin was expressed in inclusion bodies, purified and crystallized as described previously^{15,16}. The two heavy-atom derivatives were 10 mM K₂PtCl₄ (3-day soak) and 1 mM platinum(II)(2,2':6',2''-terpyridine) chloride dihydrate (Pt-terpyridine) (2-week soak). Both derivatives were made by transferring crystals into 30% saturated Li₂SO₄. Native and PtCl₄ data were collected at the Hamburg synchrotron (wavelength 0.927 Å) to approximately 3.0 Å. Pt-terpyridine data were collected to 3.5 Å at the Daresbury synchrotron, station 9.5 (wavelength 0.89 Å). All data collection was performed at a temperature of 100 K. Single crystals were used for each of the native and derivative data sets. The cryoprotectant was 35% glycerol and 65% well-solution equivalent. Data were processed using DENZO¹⁷. The resolution limit of each merged data-set was determined using an average $I/\sigma(I) > 2.0$ criterion rather than merging R -values. Shelx-90 (ref. 18) was used to solve the PtCl₄ difference Patterson, and cross-phased difference fouriers¹⁹ were used to find the 3Pt-terpyridine sites and additional PtCl₄ sites (3 in total). The heavy-atom parameters were refined using MLPHARE²⁰. The average figure of merit after refinement was rather low (0.32), corresponding to a mean phase error of about 70°. Density modification (solvent flattening and histogram matching) was performed using DM^{21,22}. The map was considerably improved after the density-modification phases were recycled and the molecular envelope refined. The resulting map (with a corresponding R_{free} of 32%) was of sufficient quality to be readily interpreted. O²³ was used to build the structure, which consists of 539 residues and 75 water molecules. The structure was refined with X-PLOR²⁴ using the R_{free} to guide the refinement. The results of any cycle of refinement were not accepted if R_{free} did not fall. After overall anisotropic temperature-factor refinement and restrained temperature-factor refinement, the conventional crystallographic residual (using all data with $F > 2\sigma$ from 6 to 2.5 Å) is 21.0% and the R_{free} is 28.0%. Of the 31,873 reflections, 29,851 make the σ -cut-off. No observed non-glycine residues are in disallowed regions of the Ramachandran plot^{25,26}. The map calculated using DM phases and f.o.m.s is shown in Fig. 2.

* $R_{\text{merge}} = \sum_h \sum_j |I(h) - \langle I(h) \rangle| / \sum_h \sum_j I(h)$, where $I(h)$ is the mean intensity for reflection h , after rejection of outliers.

† Phasing power = $\sum_h |F_{\text{Hcalc}}| / \sum_h \epsilon$, where $|F_{\text{Hcalc}}|$ is the calculated structure factor amplitude for the heavy atom, and ϵ is the residual lack of closure (of the vector triangle $|F_P|$, $|F_{\text{PH}}|$ and $|F_{\text{Hcalc}}|$) for reflection h (calculated for acentric and centric reflections over the resolution interval 30–2.4 Å).

‡ $R_{\text{cull}} = \sum_h |F_{\text{PH}} - F_P| - |F_{\text{Hcalc}}| / \sum_h |F_{\text{PH}}| - |F_P|$, where $|F_{\text{PH}}|$, $|F_P|$ and $|F_{\text{Hcalc}}|$ are defined above (summed over centric reflections). Figure of merit (f.o.m.) = $\langle \cos(\Delta\alpha_h) \rangle$, where $\Delta\alpha_h$ is the error in phase angle for reflection h .

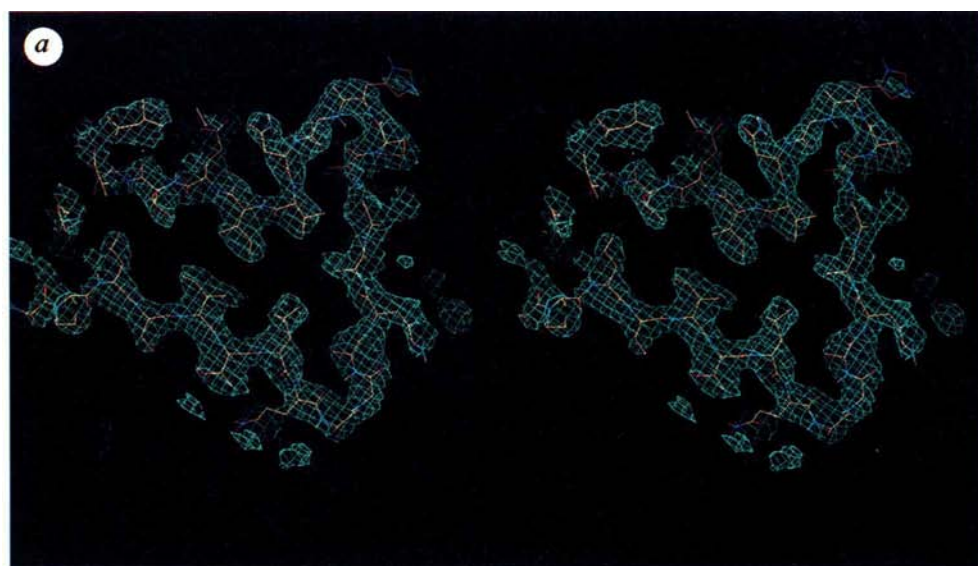
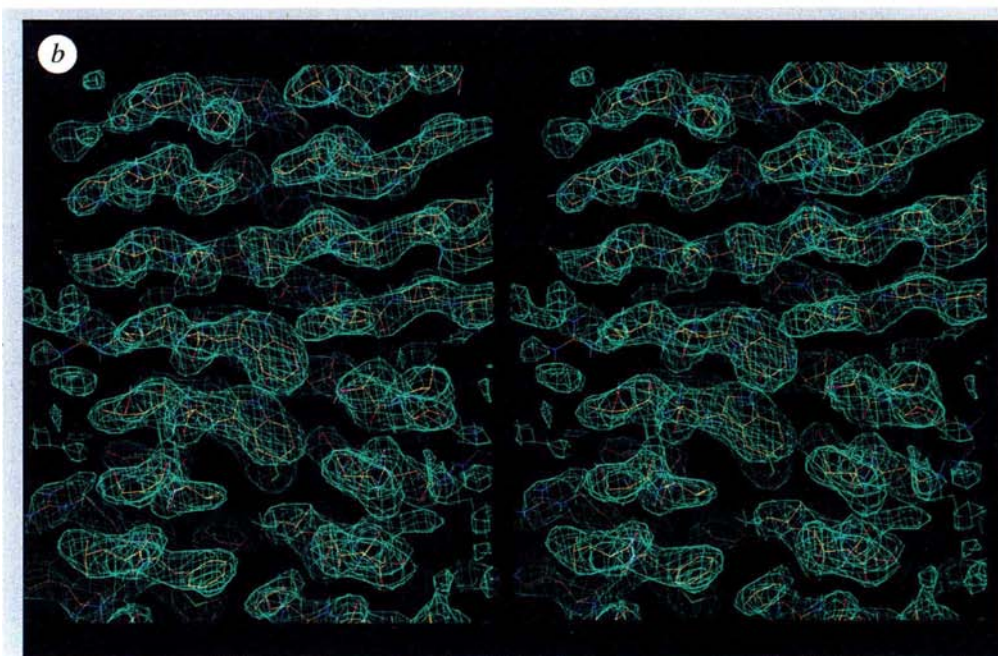


FIG. 2 Examples of DM map. Typical cross-sections of P.69 pertactin, perpendicular (a) and parallel (b) to the helical axis. The map was calculated using phases and f.o.m.s from DM (no model phases used). The model has had one round of refinement.



allow the design of experiments to increase our understanding of the mechanisms of virulence of *B. pertussis*, and also of those regions of the molecule involved in the generation of protective immunity. □

Received 10 October 1995; accepted 6 March 1996.

1. Ad Hoc group for the study of the pertussis vaccines. *Lancet* **8592**, 955–960 (1988).
2. Greco, D. et al. *New Engl. J. Med.* **334**, 341–348 (1996).
3. Gustafsson, L., Hallander, H. O., Olin, P., Reizenstein, E. & Storsaeter, L. *New Engl. J. Med.* **334**, 349–355 (1996).
4. Charles, I. G. et al. *Proc. natn. Acad. Sci. U.S.A.* **86**, 3554–3558 (1989).
5. Montaraz, J. A., Novotny, P. & Ivanyi, J. *Infect. Immun.* **47**, 744–751 (1985).
6. Charles, I. G. et al. *Microbiology* **140**, 3301–3308 (1994).
7. Hynes, R. O. *Cell* **48**, 547–554 (1987).
8. Ishibashi, Y., Claus, S. & Reiman, D. A. *J. exp. Med.* **180**, 1225–1233 (1994).
9. Leininger, E. et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 345–349 (1991).
10. Williamson, M. P. *Biochemistry* **297**, 249–260 (1994).
11. Charles, I. G. et al. *Eur. J. Immun.* **21**, 1147–1153 (1991).
12. Yoder, M. D., Lietzke, S. E. & Jurnak, F. *Structure* **1**, 241–251 (1993).
13. Baumann, U., Flaherty, K. M. & McKay, D. B. *EMBO J.* **12**, 3357–3364 (1993).
14. Yoder, M. D., Keen, N. T. & Jurnak, F. *Science* **260**, 1503–1507 (1993).
15. Makoff, A. J., Ozer, M. D., Ballantine, S. P., Fairweather, N. F. & Charles, I. G. *Biotechnology* **8**, 1030–1033 (1990).

16. Emsley, P., McDermott, G., Charles, I. G., Fairweather, N. F. & Isaacs, N. W. *J. molec. Biol.* **235**, 772–773 (1994).
17. Otwinowski, Z. in *Data Collection and Processing* (eds Sawyer, L., Isaacs, N. W. & Bailey, S.) 56–62 (Science and Engineering Research Council, Daresbury Laboratory, Warrington, UK, 1992).
18. Sheldrick, G. M. *Acta crystallogr. A* **46**, 467–473 (1990).
19. Collaborative Computing Project No. 4. *Acta crystallogr. D* **50**, 760–766 (1994).
20. Otwinowski, Z. in *Proceedings of the CCP4 Study Weekend 25–26 January 1991* (eds Wolf, W., Evans, P. R. & Leslie, A. G. W.) 80–85 (Science and Engineering Research Council, Daresbury Laboratory, Warrington, UK, 1991).
21. Wang, B. C. *Meth. Enzym.* **115**, 90–112 (1985).
22. Cowtan, K. in *Joint CCP4 and ESF-EACBM Newslett. Protein Crystallogr.* **31**, 34–38 (1994).
23. Jones, T. A., Cowan, S., Zou, J.-Y. & Kjeldgaard, M. *Acta. crystallogr. A* **47**, 110–119 (1991).
24. Brünger, A. T. *X-PLOR Reference Manual 2.1*. (Yale University, New Haven, CT, 1988).
25. Ramakrishnan, C. & Ramachandran, G. N. *Biophys. J.* **5**, 909–933 (1965).
26. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. *J. appl. Crystallogr.* **26**, 283–291 (1993).
27. Steinbacher, S. et al. *Science* **265**, 383–385 (1994).
28. Kraulis, P. J. *J. appl. Crystallogr.* **24**, 946–950 (1991).

ACKNOWLEDGEMENTS. We thank P. Novotný, M. Roberts, G. Dougan, F. Jurnak and colleagues for discussions; G. McDermott for help with data collection; and Z. Otwinowski for his opinion of the heavy-atom refinement. This work was supported by the Wellcome Trust.

CORRESPONDENCE AND MATERIALS. Requests to be addressed to N.W.I. (e-mail address N.Isaacs@Chem.gla.ac.uk).

An immunology inventory

This feature on immunology and antibodies looks at various useful products, such as an optical verification plate, an automated immunostaining system, a new free-PSA ELISA and *in situ* hybridization kits to specific cytokines.

DYNATECH has launched a new **optical plate system** to verify its microplate readers and to assist in meeting quality assurance requirements. The optical verification plate is available in either a 12-way format for the MRX family or an 8-way format for the DIAS automated EIA processor. Unlike the self-tests found on most modern instruments, which verify only that various components and assemblies are working, the optical test plate checks that the actual performance of the instrument is satisfactory. These verification plates are designed to test for alignment, linearity, channel matching, long-term drift, filter performance and reading accuracy. The plates contain neither moving parts nor reagents or dyes. Moreover, they are individually calibrated and are supplied with a certificate of calibration. A recalibration service is also offered by the company to ensure a long working life.

Reader Enquiry No. 100

Becton Dickinson Immunocytometry Systems has announced the availability of six new research **reagents specifically developed for HIV and T-cell biology studies**. They are CD4 v4 (clone L120), CD26, CD27, CD19 (clone SJ25C1) and CD34. The CD4 v4 (clone L120), CD26 and CD27 antibodies offer distinct benefits to researchers working in areas ranging from



HIV and T-cell biology reagents for high viral titres in co-culture studies — see Becton Dickinson.

high viral titres in co-culture to studies of cell signalling and naïve T lymphocytes. The CD19 (clone SJ25C1) and CD34 are now available as PerCP conjugates. These tools should prove useful for AIDS and T-cell research. CD4 v4 (clone L120) recognizes CD4 at a location discrete from the HIV binding site. In cases where

altered CD4 v1 domains may be present, these FITC and PE reagents are valuable tools for assessing expression and enumeration of CD4. Due to the fact that CD26⁺CD4⁺ T cells are preferentially infected *in vitro* and are the main reservoir for HIV-1 in infected individuals, CD26 FITC and PE reagents should be important in evaluating HIV infection levels in CD4 subsets. Cell-signalling studies may also be possible because CD26 is expressed on both T and B lymphocytes. CD27 and its ligand, CD70, are implicated in T and B signalling and the resulting T-cell activation and response.

Reader Enquiry No. 101

Automated systems

The Anthos Labtec 2010 **microplate reader** extends the company's range of instruments, combining today's measurement technology with the flexibility of a modular system. The instrument can be operated as a stand-alone instrument or linked to a PC. The stand-alone system includes an integrated palmtop PC with a system control and evaluation program, similar to the one used in the Anthos HT reader line. Qualitative, quantitative, semi-quantitative and combined evaluation, together with the validation options, enable the user to define the procedure required for a particular microplate application. A sequence of menus is designed to reduce measuring and evaluating to just a few keystrokes. The second version of the instrument, without integrated palmtop computer, can be linked to a PC and operated through the same software used with the palmtop version.

Reader Enquiry No. 102

The new **robotic manipulator arm** from Tecan, RoMa, extends the use of automation using the Genesis series of robotic

sample processors (RSP). Fitted as an extra arm in the instrument, it can transport any object with the outer dimensions of a microplate, for example deep-well plates, microplate lids and reagent blocks. It is designed to provide a further level of automation, allowing integration with other devices, such as plate washers and



Tecan's RoMa robotic manipulation arm.

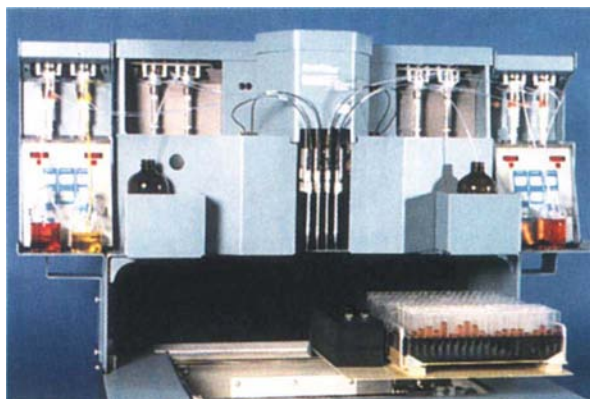
readers. A shelf can be fitted above the series' optional positive identification system, providing storage for microplates, deep-well plates and reagents. The RoMa arm is available as an option with new RSP series instruments, or can be retrofitted to existing installations. This series with robotic manipulator arm can operate with a wide range of manipulative robots, including Tecan's Trac.

Reader Enquiry No. 103

Zymed and Leica have announced that they are jointly offering the ST5050 **automated immunostaining system**, which includes the ST5050 immunostainer and HistoST5050 reagents. This flexible and reproducible immunostaining system is designed for general immunohistochemistry applications and for routine laboratory use. The ST5050 is a second-generation automated immunostaining instrument based on the Jung Histostainer Ig technology. Up to ten programs may be run simultaneously under temperature-controlled conditions from ambient to 45 °C. All programs and run parameters are documented to help meet the requirements set by good laboratory practice working procedures. Zymed's HistoST5050 reagents include an extensive line of primary antibodies (tumour and disease markers), and detection kits specially optimized for use with the immunostainer. These reagents utilize second-generation, streptavidin-biotin labelling technology and an advanced single component substrate system, which is said to result in enhanced sensitivity with low background. Operator handling should be minimized, as all the reagents are ready-to-use, single-solution components.

Reader Enquiry No. 104

ICN Pharmaceuticals has developed the Titertek Quadflex, a flexible **automated liquid-handling device** with a wide range of applications for research and clinical use. Integrated with the Titertek S20 stacker and microplate reader, the system is said to be a complete EIA processing system. It uses four spreadable, non-metallic, liquid-level sensing tips and robotic *x/y* moving platform to address user-defineable sample and receiver vessels. This platform, in conjunction with the 100 per cent glass or Teflon liquid path, is said to ensure accurate, precise



ICN Pharmaceuticals' complete EIA processing system.

liquid delivery in volumes as low as 2 μ l, with minimal carryover. The software allows for complete user control over all aspects of the assays, including sample and reagent volumes/locations, wash and incubation cycles, and reading options. The manifold is said to allow rapid, high-throughput processing of microplates and production of 'mother-daughter' plates for drug-screening applications.

Reader Enquiry No. 105

ELISAs

Amersham has launched a range of nine **human cytokine ELISAs**, now available as an addition to the Biotrak cellular communication assays. This should enable

researchers to measure cytokines at 'normal' concentrations (as low as 0.1 pg ml^{-1}) in serum/plasma samples using no more protocol steps than a conventional ELISA. The Amdex conjugate is a preformed, highly stable complex containing many hundreds of horseradish peroxidase (HRP) molecules instead of the usual monomeric HRP conjugate. The high number of enzyme molecules enables the rapid generation of an amplified signal. The ELISAs are available to measure IL-1 α , IL-1 β , IL-4, IL-6, IL-10, GM-CSF, IFN α , IFN γ and TNF α . Each assay kit enables the measurement of 41 unknowns in duplicate and includes extra diluent, allowing samples and standards to be diluted with the same reagent.

Reader Enquiry No. 106

Endogen has released **human matched antibody pairs and recombinant standards** for use in ELISAs. The matched antibody pairs and cytokine standards provide researchers with reagents to develop cost-effective cytokine ELISAs in their own laboratory. Coating

and detecting antibodies have been validated specifically for ELISA applications. Detection antibodies are biotinylated, and can be used with a variety of streptavidin detection systems. Antibodies are provided as concentrates and cytokine standards are lyophilized, including hIL-1 β , hIL-4, hIL-6, hIL-10, hIFN γ and hTNF α .

Reader Enquiry No. 107

With five Cytoscreen **rat-specific ELISA** kits available, BioSource offers a broad range of rat kits pertinent to today's research applications, such as AIDS, lupus, rheumatoid arthritis, hepatitis, cachexia and protozoan infections. This ELISA kit line features rat IFN- γ , MCP-1, MIP-2, TNF- α , and TNF- α UltraSensitive. Incubation

times with this line of ELISAs range from 2 h 30 min to 4 h 50 min, depending on the kit. Sensitivity ranges from <1 pg ml^{-1} to <8 pg ml^{-1} ; the TNF- α UltraSensitive ELISA kit has a sensitivity of <0.7 pg ml^{-1} . All the rat kits have a stated precision of <9 per cent inter-assay and <5 per cent intra-assay. BioSource offers 47 ELISA kits for human, rat, mouse and monkey cytokines, chemokines and interferons. Cytoscreen kits contain all the necessary reagents for 192 determinations.

Reader Enquiry No. 108

CytImmune Sciences has announced the availability of the CytELISA *in vitro* **cytokine enzyme immunoassay kits** for the detection of human IL-1 β , IL-2, IL-4, IL-6, IL-10, TNF α and VEGF in cell culture supernatants. The kits are sandwich-format assays designed to measure cytokines in cell-culture supernatants. A neutralizing monoclonal antibody capture system is designed to ensure the detection of 'free' cytokine. The formazin-colour amplification system is designed to provide a high signal-to-noise ratio and increased sensitivity. The five-step protocol can be run in 2.5 h and it has a dynamic assay range of 8.0–2,000 pg ml^{-1} of cytokine, which means that concentrating the sample is not necessary. Two 96-well plates and sufficient reagents for 192 determinations are provided.

Reader Enquiry No. 109

Antibodies

Maine Biotechnology now offers **monoclonal and polyclonal antibodies to angiotensin factor-1 (AT1) receptor**. These antibodies have specificity against naturally occurring antigen receptors by western blot and react positively by ELISA to cells that express the AT1 receptors. Two major subtypes of angiotensin II receptors have been cloned and are known as the AT1 and AT2 receptors. The AT1 receptor appears to mediate all of the vasoactive properties of angiotensin II. The renin-angiotensin system has a broad tissue distribution and also plays a role in salt and water metabolism. Future applications for these products may include clinical diagnostic assay development.

Reader Enquiry No. 110

Four new **human chorionic gonadotrophin (hCG) monoclonal antibodies** from Genzyme constitute a new line for the company, designed to enhance performance in applications ranging from immunochromatography, through ELISA and IFA, to immunohistochemistry and western blots. These products include MIH9834, MIH9836, MIH9837 and MIH9838. The first is MIH9834, which shows high reactivity with intact hCG and β -hCG and possibly eliminates the effects of the LH scavenger.

ADVERTISEMENTS



Gramsch Laboratories

- New: Somatostat. Rec. 2A/B Antisera
- G-Protein Antisera - α , β , γ -subunit
- G-Protein Peptide Fragments
- Opioid Receptor Antisera
- Monoclonal Opioid-Antibody 3E7 YGGF... - specific; appl.: combinatorial peptide library tests
- Cyclic Nucleotide Antisera
- Antibody/Peptide Production Service

Phone 08138 1769 or 08131 5679-0
Fax 08138 1260 or 08131 5679-24
WWW: ourworld.com/serve.com/
homepages/jgramsch/gf.htm

Gramsch GmbH

D-85247 SCHWABHAUSEN, Germany

NEWCASTLE UNIVERSITY
MOLECULAR BIOLOGY UNIT



PROTEIN SEQUENCING

Membrane-bound, solid or solution samples

£12 per cycle

5 cycles minimum - no set up charge

For further information contact Dr. Joe Gray

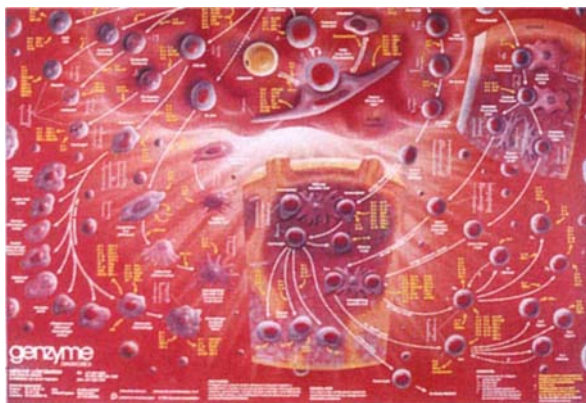
Phone: +44 191 222 8612

Fax: +44 191 222 8129

Email: JOE.GRAY@newcastle.ac.uk

READER ENQUIRY NO. 16

READER ENQUIRY NO. 12



Genzyme also offers educational slides to help illuminate the cytokine pathways in the immune system.

Useful in immunochromatographic applications, MIH9834 also pairs well with the second antibody, MIH9836. MIH9836 has highly specific binding only with intact hCG, not α - or β -hCG. This makes it useful in a double monoclonal sandwich assay for intact hCG. The third antibody, MIH9837, shows low reactivity with other glycoprotein hormones, reacting predominantly with intact hCG. It is therefore suitable for quantification of intact hCG in sandwich immunoassays. Finally, MIH9838 is specific to the α subunit of hCG and other glycoprotein hormones. It also pairs well with the other three antibodies. Full pairing information is available and all the products are available in bulk volumes for convenience. The company also offers educational slides covering cytokine pathways in the immune system.

Reader Enquiry No. 111

New from R&D Systems is a range of *in situ* hybridization (ISH) kits specific for particular cytokines. They permit detection of human IL-6, IL-10, TGF- β 1 or TNF- α mRNA. Kits are also available for the detection of murine IL-10 or rat/murine TGF- β 1 mRNA. Target mRNA is detected using a cocktail of seven labelled oligonucleotide probes and is then visualized using an antibody conjugated to alkaline phosphatase. This technique does not make use of radiolabelled probes. Each kit contains a set of positive and negative controls. Positive control tissue slides are also supplied. All the kits have been fully validated using a number of cell lines and are suitable for use with paraffin-embedded sections, frozen-tissue sections and cell smears.

Reader Enquiry No. 112

BioGenex offers the monoclonal antibody 9E10 that reacts with an epitope of c-myc protein, which may be detected in many primary cancers, including breast, lung, colorectal, cervical, liver, stomach, thyroid, brain, kidney and prostate, as well as leukaemia, lymphomas and melanomas. The c-myc proto-oncogene was initially

identified through study of the cancer-inducing oncogene present in the MC29 avian retrovirus. Activated proto-oncogenes play an important role in the development and progression of human cancers. Monoclonal antibody 9E10 reacts specifically with c-myc protein in formalin-fixed and paraffin-embedded tissue sections. This antibody is available in concentrated form or as a ready-to-use reagent, optimized for use with the Super Sensitive biotin-streptavidin detection systems. The antibody

is currently available for research use only.

Reader Enquiry No. 113

Diagnostics

Shield Diagnostics has launched a new immunoassay for ristocetin co-factor to detect the von Willebrand factor (vWF) activity. vWF disease is one of the most common inherited bleeding disorders, with a known incidence of 1:8,000. Diagnosis of vWF disease requires detection of von Willebrand factor antigen and vWF activity. Both tests are required to differentiate between subtypes of the disease and identify appropriate therapeutic modalities for intervention. The new ELISA-based assay for the determination of vWF activity provides a sensitive alternative to the standard platelet aggregation methods. The vWF activity ELISA contains 12 $\times$ 8 well microtitre plate strips coated with a monoclonal antibody, which recognizes the functional epitope of vWF. The kits include anti-vWF antibody conjugated to horseradish peroxidase, which is used to detect vWF bound to the monoclonal antibodies, as well as all controls, calibrators and reagents required. According to the manufacturer, the ELISA has an inter-assay coef-

ficient of variation of eight per cent and an intra-assay coefficient of variation of four per cent.

Reader Enquiry No. 114

Immunotech now distributes the Coulter HIV-1 p24 antigen assay and the HIV-1 p24 antigen neutralization kit for *in vitro* diagnostic use. This assay can be used for qualitative and quantitative detection of HIV-1 p24 in human serum, plasma or tissue culture supernatants. This product is intended to be used for screening individuals at unknown risk for HIV-1 infection



The Coulter HIV-1 p24 antigen assay.

and as an aid in diagnosis, prognosis and monitoring of HIV-1 infection and progression. According to the manufacturer, the immunoassay takes three hours, and is supplied in 96- and 2,400-assay sizes.

Reader Enquiry No. 115

An ELISA kit to monitor chemotherapeutic agents in cancer research has been introduced by BioWhittaker. The Matri-Tect NMP41 cell death test kit may be used to quantitate human cell death and may also find use in the study of Alzheimer's disease and organ transplant rejection. The kit takes advantage of a recent technique in the study of malignancy and disease progression, detecting specific nuclear matrix protein (NMP) known as nuclear mitotic apparatus protein (NuMA). As a result of cell death, soluble NMPs are released by dead and dying cells and may be detected in serum

ADVERTISEMENTS

Now available in ton lots from EU production:

Di-*tert*-butyldicarbonate

Protecting agent for peptide synthesis

CC(C)(C)OC(=O)OC(=O)OC(C)(C)C

Distributed by:

Produced by:
CHEMTEC LEUNA
Gesellschaft für Chemie und Technologie mbH

SIBER HEGNER RAW MATERIALS LTD.
Wiesenstrasse 8, P.O. Box 888 Phone: +41-1-386 73 36
CH-8034 Zurich, Switzerland Fax: +41-1-383 16 63

DNA SEQUENCING

£45 per reaction

- PCR products
- single strand DNA
- M13 inserts
- double strand DNA
- plasmid inserts
- no result? no charge!

By research labs for research labs

Molecular Biology Services
King's College School of Medicine and Dentistry
Rayne Institute, 123 Coldharbour Lane, London SE5 9NU
Tel. 0171-346 3126 Fax: 0171-733 3877

READER ENQUIRY NO. 2

READER ENQUIRY NO. 11

plasma or cell culture supernatants. By measuring NMP levels after administration of a chemotherapeutic agent it is possible to measure the effectiveness of the agent and deduce its suitability for use in humans. As a result of cell death in the brain, measurement of spinal fluid may show an increase in the levels of NMP41 present. The kit may also prove useful in the study of organ transplant rejection, organ failure and tissue death, which result in a rise in NMP41 levels. Detection of these NMP41 levels could provide rapid and accurate patient assessment.

Reader Enquiry No. 116

IDS now offers CanAg Diagnostics' new **free-PSA ELISA**, to give increased information in the diagnosis and management of prostate cancer. Determination of total PSA in serum is used for the detection and management of patients with prostate cancer. However, there can be large variations in the relationship between free PSA and the PSA-ACT complex (total PSA) in different individuals. Recent research suggests that measuring both forms of the antigen, and calculating the ratio of free-to-total forms can help make a clearer distinction between prostate cancer and benign prostatic hyperplasia. This ELISA has a 12 × 8 well microtitre format and short incubation times, producing results in about 2.5 hours. The assay is sensitive, measuring as low as 0.05 µl l⁻¹ using a 50-µl sample. The kit is designed to be used with the company's Equi-Molar PSA ELISA, which is said to show identical specificity for both free and complexed PSA, to give a true measurement of total PSA levels.

Reader Enquiry No. 117

Literature

Affiniti Research Products offers Transduction Laboratories' range of high-quality **antibodies for signal transduction research**. The newly released 1996 catalogue details some 250 monoclonal and polyclonal antibodies of high quality and of broad applicability. All antibodies are supplied with positive-control cell lysates and detailed data and application sheets. The catalogue includes full details of each antibody, their derivation, their utility and their specific species reactivity. Additionally, detailed applications protocols, published reference lists and a trouble-shooting guide are provided.

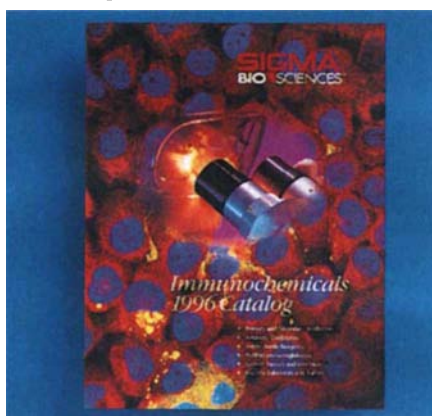
Reader Enquiry No. 118

The new 1996 Serotec **product guide** has 128 pages and features over 2,500 products, including new human CD specificities. These include RPE-Cy5 conjugates for three-colour, flow-cytometry staining and new preservative-free adhesion markers provided for functional studies of humans, mice and rats. There is a full range of growth factor and cytokine reagents, over

75 cell-signalling antibodies and new hormone receptor antibodies. Also included are new formats of inhibin and pro-alpha C assay kits, acetylcholine receptor antibodies, new anti-mouse cell surface markers, and anti-rat cell surface markers.

Reader Enquiry No. 119

The 1996 **immunochemicals catalogue** now available from Sigma Chemical contains details of 1,700 immunochemicals. Featuring 130 new products, as well as a new 'clone index' for quickly locating monoclonal products, this 484-page desktop reference includes a comprehensive selection of primary and secondary antibodies, antibody conjugates, avidin/biotin reagents, purified immunoglobulins, growth factors, interleukins, enzyme substrates and buffers. The catalogue also includes over 60 new growth factors and antibodies,



The 1996 immunochemicals catalogue.

including angiogenin, platelet-derived endothelial cell growth factor and pleiotrophin, as well as ELISA kits and a new selection of chemokines. There are also new products listed for analytical cytometry, cell and molecular biology, and neuroscience. The items listed in the catalogue are supported by toll-free technical and applications assistance.

Reader Enquiry No. 120

Odds and ends

Wako BioProducts' new **immunoblotting kit** is designed to provide lower background levels and increased sensitivity. According to the manufacturer, the immunoblotting kit enhances detection capabilities by up to two times, taking antigen detection levels to the level of radioisotope detection. The kit uses Luminol oxidation-HRP chemiluminescence detection with a thiazol enhancement, coupled with a streptavidin-biotin system (ABC reagent). The chemiluminescent reaction, provided by the combination of Luminol substrate with horseradish peroxidase targeting enzyme, is enhanced by thiazol and is said to allow for quick, strong signalling.

Reader Enquiry No. 121

Incstar has compiled a complete **listing of its immunohistochemistry research products**, which includes colour photographs of antibodies utilizing various immunolabelling techniques in tissues. The immunostaining of tissues is demonstrated via indirect immunofluorescence, peroxidase, antiperoxidase and biotin/avidin techniques. This brochure allows the customer to visualize the high-quality staining expected when utilizing the company's histochemical products. As part of its overall customer service program, Incstar also provides an immunohistochemical instruction manual, which contains useful information on immunolabelling techniques, fixative and buffer formulas, perfusion techniques, stain enhancement and troubleshooting. A technical service staff on hand to provide the assistance necessary to achieve optimal results with a variety of immunohistochemical applications.

Reader Enquiry No. 122

These notes are compiled by Brendan Horton from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal.

ADVERTISEMENTS

DNA/RNA Custom Synthesis

of high quality and purity with high and low scale for all applications

*** HPLC purification, mixed base sites, ANTISENSE, and all oligo modifications available**

*** Special Offers Through 8-31-96 and Guaranteed Low Prices**



Bahnstrasse 2 D-97450 Arnstein-Bindfeld

A NEW FRONTIER IN GENE TECHNOLOGY
Fax Ordering: +49-(0)9360-1809

READER ENQUIRY NO. 4

Custom MAb Production

- Hybridoma Development
- Ascites Production
- In-Vitro Production
- Purification Services
- Labeling • Fragmentation

(518) 537-8000



READER ENQUIRY NO. 5